Stephen M. Zemyan

The Classical Theory of Integral Equations

A Concise Treatment

Birkhäuser

Stephen M. Zemyan

The Classical Theory of Integral Equations

A Concise Treatment

Birkhäuser

Stephen M. Zemyan
Department of Mathematics
Pennsylvania State University Mont Alto
Mont Alto, PA, USA

ISBN 978-0-8176-8348-1 ISBN 978-0-8176-8349-8 (eBook)
DOI 10.1007/978-0-8176-8349-8
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2012939157

Mathematics Subject Classification (2010): 45Bxx, 45Dxx, 45Exx, 45Fxx, 45Gxx

Printed on acid-free paper

Springer is part of Springer Science+Business Media
(www.birkhauser-science.com)

Dedicated To My Parents
Florence E. Zemyan
and
Edward M. Zemyan

Preface

The study of integral equations is a thoroughly fascinating chapter in man's continuing search for mathematical understanding, and the outcome of this search is both strikingly beautiful and intrinsically interesting.

The greater part of the theory of integral equations was developed early in the twentieth century as a result of the efforts of many brilliant individuals, and most of the important integral equations fittingly bear their names.

One of the most important categories of integral equations is the Fredholm integral equation, which was named after the renowned Swedish mathematician Erik Ivar Fredholm (April 7, 1866 to August 17, 1927). His landmark paper, *Sur une classe d'equations fonctionelles*, was published in *Acta Mathematica* in 1903. The first three chapters of this text are devoted to linear Fredholm integral equations.

Another important category of integral equations is the Volterra integral equation, which was named after the distinguished Italian mathematician Vito Volterra (May 3, 1860 to October 11, 1940). Chapter 4 is devoted to linear Volterra integral equations. Nonlinear Volterra integral equations are briefly discussed in Chap. 6, and singular Volterra integral equations are touched upon in Chap. 7.

The prolific German mathematician David Hilbert (January 23, 1862 to February 14, 1943) made huge contributions to the foundation of the general theory of integral equations in his tome *Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen*. The Hilbert–Schmidt theorem in Chap. 3 and the Hilbert transform in Chap. 7 are essential tools in the field.

In little over a century, more than 11,000 articles and dozens of books and manuscripts have been written that concern various aspects of the theory of integral equations. Perhaps the best way to obtain an overview of this vast subject area would be to examine the 2010 Mathematics Subject Classification published by the American Mathematical Society. This classification appears in Appendix A for the convenience of the reader.

Although it is certain that the theory of integral equations is an important part of pure mathematics, it is also true that there are many applications of this theory to problems in the physical sciences. Furthermore, considerable interactions exist

between the area of integral equations and other areas within mathematics, such as linear algebra, operator theory, and ordinary differential equations.

The specific goal of the author of any textbook is to enrich the academic lives of students by constructing an effective tool for learning and teaching, as well as independent study. My primary intention in writing this text is to present a rigorous and systematic development of the classical theory of integral equations that is accessible to the advanced undergraduate or early graduate student. Since most students at this level and many practicing scientists are generally not familiar with the intricacies of Lebesgue integration or the theory of integral operators at a sophisticated level, no references are made to these theories here. Yet, it is still possible to present most of the main results by assuming only that the functions that appear in an integral equation are either continuous or integrable in the sense of Riemann. It is also possible to give a rather thorough treatment of many significant theorems involving integral operators without the intense sophistication required in an advanced course of study. Indeed, much of the theory was originally derived under these relaxed assumptions. Hopefully, our presentation will serve as a firm foundation for the theory as well as a springboard for further reading, exploration, and research.

Although no previous experience with the theory of integral equations is required, some prerequisites are essential. It is assumed that the reader has some expertise in reading as well as writing mathematics. It is also assumed that the reader is generally familiar with the important definitions and theorems in the subject areas of linear algebra and advanced calculus that are necessary to understand the development of the theory here. Undoubtedly, a concise review of the prerequisites would refresh the student's memory. In order to serve this purpose and to enhance the overall completeness of the presentation, a section entitled *Tools of the Trade* appears at the beginning of each chapter.

In preparing this text, I have striven for precision and completeness in every example, explanation, or exercise. Theorems and their proofs are presented in a rigorously analytical manner. No tricky, frustrating details are "easily verified" or "left to the reader." Furthermore, the reader is not asked to refer to other texts to search for results that are supportive of statements made in this text. Hopefully, the result is a self-contained, straightforward, and crystal-clear treatment of the theory.

There is considerable and justified emphasis on the methods of solution. Since it might be extremely difficult or even impossible to obtain an exact solution to an integral equation, numerical methods for computing approximate solutions assume great importance. Comprehensive examples are presented that reinforce underlying theories and illustrate computational procedures. When a particular numerical procedure is especially tricky or complicated, a chart entitled *A Concise Guide to Computation* is inserted into the text.

The problem sets are meant to be devices for enrichment. Since mathematics is learned by doing, not just reading and memorizing results, understanding is generally enhanced by struggling with a variety of problems, whether routine or challenging. Understanding may even be enhanced when a problem is not solved completely. Some problems are intended to reinforce the theory while others are

purely computational. A problem may also introduce new material or relate the material of that section to other areas of mathematics.

Some expertise in the use of Wolfram's *Mathematica* or a similar software package is assumed so that students can solve problems similar to those presented in the illustrative examples, verify solutions to the exercises that require computational intensity, and engage in creative investigations of their own.

In conclusion, I have endeavored to construct a completely comprehensible exposition of the theory, so that students who read it will be adequately prepared to solve a wide variety of problems. My hope is that the book will provide the background and insight necessary to facilitate a thorough understanding of the fundamental results in the classical theory of integral equations, as well as to provide the motivation for further investigation.

Mont Alto, PA Stephen M. Zemyan

Acknowledgements

No manuscript of this magnitude is possible without the assistance and support of others.

First and foremost, I would like to express my everlasting appreciation to my parents, Edward and Florence Zemyan, for instilling in me the value and importance of education, hard work, and persistence. I am infinitely indebted to them, not only for the academic skills that have enabled me to write this textbook, but also for lifelong support in other areas of my life.

Jack Ryff read draft copies of each chapter, spotting typographical errors, checking calculations and proofs, and offering suggestions for improvement that ultimately enhanced the text.

My employer, The Pennsylvania State University, has provided long-term academic support in the form of library services in addition to the requisite hardware and software.

Hearty thanks are owed to my editor, Tom Grasso, and his assistant Katherine Ghezzi, for their guidance and continued assistance. Their assistants and the production staff at Birkhäuser Boston made the entire publishing process as smooth, efficient, and painless as possible. The external reviewers whom I asked to examine my work provided many helpful comments regarding its content and exposition.

The numerous calculations necessary to present the illustrative examples and exercises throughout the text were made with Wolfram's *Mathematica 8.0* or previous versions of it.

I also wish to acknowledge the American Mathematical Society for granting permission to include the section of the 2010 Mathematics Subject Classification pertaining to integral equations in Appendix A. This section allows students to get an immediate overview of the organization of the subject.

Finally, portions of three previously published articles were reproduced in Appendix B with the kind permission of Springer Science+Business Media B.V. These three cited articles are available in their entirety on the SpringerLink web site.

Contents

1 Fredholm Integral Equations of the Second Kind (Separable Kernel) 1
1.1 Tools of the Trade 2
1.2 An Illustrative Example 4
1.3 The Fredholm Theorems 13

2 Fredholm Integral Equations of the Second Kind (General Kernel) ... 31
2.1 Tools of the Trade 31
2.2 The Method of Successive Substitution 33
2.3 The Method of Successive Approximation 37
2.4 The Fredholm Theorems 48
2.5 Constructing the Resolvent Kernel 60
2.6 Numerical Methods 69
2.6.1 The Method of Uniform Approximation 69
2.6.2 The Method of Collocation 72
2.6.3 Quadrature Methods 78

3 Fredholm Integral Equations of the Second Kind (Hermitian Kernel) 85
3.1 Tools of the Trade 85
3.2 Hermitian Kernels and Operators 95
3.2.1 Properties of Hermitian Kernels and Operators 97
3.2.2 The Eigenvalues of a Hermitian Kernel 98
3.2.3 The Eigenfunctions of a Hermitian Kernel 105
3.3 Finite Bilinear Expansions 110
3.4 The Hilbert–Schmidt Theorem 120
3.5 Numerical Methods 140
3.5.1 The Method of Traces 140
3.5.2 The Method of Rayleigh and Ritz 142

4 Volterra Integral Equations 151
4.1 Tools of the Trade 152

4.2 The Method of Successive Approximation 153
4.3 Convolution Kernels 166
4.4 Equations of the First Kind 172
4.5 Numerical Methods 180

5 Differential and Integrodifferential Equations 183
5.1 Tools of the Trade 183
5.2 Boundary Value Problems 189
5.2.1 Conversion to a Fredholm Integral Equation 189
5.2.2 Conversion from a Fredholm Integral Equation 191
5.3 Initial Value Problems 197
5.3.1 Conversion to a Volterra Integral Equation 197
5.3.2 Conversion from a Volterra Integral Equation 198
5.4 Integrodifferential Equations 205

6 Nonlinear Integral Equations 211
6.1 Tools of the Trade 213
6.2 Existence and Uniqueness Theorems 216
6.2.1 Nonlinear Fredholm Integral Equations 216
6.2.2 Nonlinear Volterra Integral Equations 217
6.2.3 Hammerstein's Equation 219
6.3 Illustrative Examples 221

7 Singular Integral Equations 243
7.1 Tools of the Trade 244
7.2 Equations with Continuous Kernels 251
7.2.1 Equations Solved with the Mellin Transform 253
7.2.2 Equations Solved with Fourier Transforms 256
7.2.3 Equations of Volterra Type 262
7.3 Equations with Singular Kernels 267
7.3.1 Abel's Equation 267
7.3.2 Equations with Cauchy and Hilbert Kernels 273

8 Linear Systems of Integral Equations 287
8.1 Tools of the Trade 288
8.2 Systems of Volterra Equations 290
8.3 Differential Equations 296
8.4 Systems of Fredholm Equations 300
8.5 Systems of Singular Equations 307

A The 2010 Mathematics Subject Classification for Integral Equations . 311

B Specialized Vocabularies and Sample Translations 315
B.1 French 316
B.1.1 Specialized French Vocabulary 316
B.1.2 Sample Translation from French into English 318

B.2 Italian 323
B.2.1 Specialized Italian Vocabulary 323
B.2.2 Sample Translation from Italian into English 326
B.3 German 330
B.3.1 Specialized German Vocabulary 330
B.3.2 Sample Translation from German into English 333

References 337

Index 341

Chapter 1
Fredholm Integral Equations of the Second Kind (Separable Kernel)

In this chapter, our purpose is to examine the Fredholm integral equation of the second kind

$$\phi(x) = f(x) + \lambda \int_a^b K(x,t)\,\phi(t)\,\mathrm{d}t,$$

where $K(x,t)$ is a separable[1] kernel.

In our analysis, we make the following assumptions:

- The unknown function $\phi(x)$ is always assumed to be integrable in the sense of Riemann so that the integral equation itself makes sense.
- The *free term* $f(x)$ is assumed to be complex-valued and continuous on the interval $[a,b]$. If $f(x) \equiv 0$ on the interval $[a,b]$, then the integral equation is called *homogeneous*; otherwise, it is called *inhomogeneous*.
- The complex constant λ $(\neq 0)$ is a parameter that should not be absorbed into the kernel.
- The *kernel* $K(x,t)$ is assumed to be complex-valued and continuous on the square $Q(a,b) = \{(x,t)\colon a \le x \le b, a \le t \le b\}$.

A kernel $K(x,t)$ is called *separable* if it assumes the specific form

$$K(x,t) = \sum_{i=1}^{n} a_i(x)\,\overline{b_i(t)}.$$

We will always assume that each function $a_i(x)$ and $b_i(t)$ is complex-valued and continuous on the interval $[a,b]$. We also assume without any loss of generality that the sets $\{a_i(x)\}$ and $\{b_i(t)\}$ are linearly independent on the interval $[a,b]$; otherwise, we could diminish the number of addends in the sum.

[1] Separable kernels are also known as *degenerate* kernels or, in older texts, as *Pincherle–Goursat* kernels.

S.M. Zemyan, *The Classical Theory of Integral Equations: A Concise Treatment*,
DOI 10.1007/978-0-8176-8349-8_1, © Springer Science+Business Media, LLC 2012

We will also explain the significant role played by the *adjoint equation*

$$\psi(x) = \overline{\lambda} \int_a^b \overline{K(t,x)}\, \psi(t)\, \mathrm{d}t$$

in the overall process. The reason for considering these equations in tandem will become apparent in the exposition.

In Sect. 1.1, we present several tools of the trade that are indispensible for the comprehension of the material in this chapter. It will shortly become evident to the reader that any thorough treatment of the theory of linear integral equations is critically dependent upon the theory of linear algebra. Consequently, we provide here a concise review of the essential information from that discipline that will used within this chapter.

In Sect. 1.2, we examine a specific Fredholm integral equation to illuminate by example the theory that supports the techniques used to determine its solution and to motivate the discussion of the theory to follow.

In Sect. 1.3, we prove the four Fredholm theorems and the Fredholm Alternative Theorem for Fredholm integral equations of the second kind with a separable kernel.

1.1 Tools of the Trade

In this chapter, the reader should be familiar with the following topics:

- *Systems of linear equations*: The theory of linear algebra is concerned with systems of linear equations and their solutions. An arbitrary system of n linear equations in n unknowns can be written in the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned}$$

  or more compactly in the form $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{A}$ is an $n \times n$ matrix and $\mathbf{x}$ and $\mathbf{b}$ are column vectors with n entries. We assume that the entries a_{ij}, x_i, and b_i are complex numbers. If $\mathbf{b} = \mathbf{0}$, then the system is called *homogeneous*; otherwise, it is called *inhomogeneous*. Every system of linear equations will have either exactly one solution, an infinite number of solutions, or no solutions at all.
- *Inverse of a matrix*: If the determinant $\det(\mathbf{A})$ of the $n \times n$ matrix $\mathbf{A}$ does not vanish, then $\mathbf{A}$ is invertible, and its *inverse* $\mathbf{A}^{-1}$ is given by

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \mathrm{adj}(\mathbf{A}),$$

where the adjoint matrix $\text{adj}(\mathbf{A})$ is defined as the transpose of the matrix of cofactors C_{ij} from $\mathbf{A}$, i.e., $\text{adj}(\mathbf{A}) = (C_{ji})$, where $C_{ij} = (-1)^{i+j} M_{ij}$ and M_{ij} is the determinant of the submatrix of $\mathbf{A}$ that remains after the ith row and jth column are deleted from $\mathbf{A}$. When $\mathbf{A}$ is invertible, the system $\mathbf{Ax} = \mathbf{b}$ has a unique solution that can be written in the form $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$. If $\det(\mathbf{A}) = 0$, then a system of linear equations will either have an infinite number of solutions or no solutions at all.

- *Eigenvalues and eigenvectors*: A scalar μ is called an *eigenvalue of the $n \times n$ matrix* $\mathbf{A}$ if there exists a nonzero vector $\mathbf{x}_\mu$ such that $\mathbf{A}\mathbf{x}_\mu = \mu\mathbf{x}_\mu$. Any solution vector $\mathbf{x}_\mu$ is called an *eigenvector of* $\mathbf{A}$ *corresponding to* μ. In other words, the eigenvectors that correspond to μ are the nontrivial solutions of the homogeneous linear system $(\mathbf{A} - \mu\mathbf{I})\mathbf{x}_\mu = \mathbf{0}$. Nontrivial solutions to this system exist if and only if $\mathbf{A} - \mu\mathbf{I}$ is not invertible, i.e., if $\det(\mathbf{A} - \mu\mathbf{I}) = 0$. There always exists at least one eigenvector corresponding to each eigenvalue. The set of all eigenvectors corresponding to the eigenvalue μ is a vector space, which we refer to as the *eigenspace* $\mathbf{E}_\mu$ *corresponding to* μ. The determinant $\det(\mathbf{A} - \mu\mathbf{I})$ is a polynomial of degree n in the variable μ, called the *characteristic polynomial of* $\mathbf{A}$. It always has n zeroes counting multiplicity by the Fundamental Theorem of Algebra. If μ is an m-fold root of the characteristic polynomial, then we say that μ has *algebraic multiplicity* m. If there exist exactly p linearly independent eigenvectors corresponding to μ, then we say that the eigenvalue μ has *geometric multiplicity* p, and $\dim(\mathbf{E}_\mu) = p$. The geometric multiplicity may be less than the algebraic multiplicity, but in any case we always have $1 \le p \le m \le n$.
- *Complex inner product spaces*: A vector space W over the field of complex numbers $\mathbb{C}$ is called a *complex vector space*. An *inner product on a complex vector space* W is a function that associates a complex number $\langle \mathbf{u}, \mathbf{v} \rangle$ with each ordered pair of vectors $\mathbf{u}$ and $\mathbf{v}$ in W in such a way that the following four axioms are satisfied for all scalars k and all vectors $\mathbf{u}, \mathbf{v}, \mathbf{w} \in W$:

  1. $\langle \mathbf{u}, \mathbf{v} \rangle = \overline{\langle \mathbf{v}, \mathbf{u} \rangle}$
  2. $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$
  3. $\langle k\mathbf{u}, \mathbf{v} \rangle = k \langle \mathbf{u}, \mathbf{v} \rangle$
  4. $\langle \mathbf{u}, \mathbf{u} \rangle \ge 0$ and $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ iff $\mathbf{u} = \mathbf{0}$

  A complex vector space with an inner product is called a *complex inner product space* or a *unitary space*.

  Two nonzero vectors are called *orthogonal* in a complex inner product space if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$. Of course, $\langle \mathbf{0}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{0} \rangle = 0$ for any $\mathbf{v} \in W$.

  The inner product is also *conjugate linear*, i.e., the properties $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$ and $\langle \mathbf{u}, k\mathbf{v} \rangle = \overline{k} \langle \mathbf{u}, \mathbf{v} \rangle$ also hold.

  One example of a complex inner product space is the vector space $\mathbb{C}^n$ that consists of all n-dimensional complex-valued vectors. Let $\mathbf{u} = (u_1, \ldots, u_n)$ and $\mathbf{v} = (v_1, \ldots, v_n)$ be any two elements of $\mathbb{C}^n$. Then the product

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{v}^* \mathbf{u} = \sum_{i=1}^{n} u_i \overline{v_i}$$

  defines an inner product on $\mathbb{C}^n$ since it satisfies all of the required axioms.

Another example is the vector space $\mathcal{C}[a,b]$ that consists of all complex-valued functions that are continuous on the interval $[a,b]$. Let $f(x)$ and $g(x)$ be any two elements of $\mathcal{C}[a,b]$. Then the product

$$\langle f,g\rangle = \int_a^b f(x)\,\overline{g(x)}\,\mathrm{d}x$$

defines an inner product on $\mathcal{C}[a,b]$ since it also satisfies all of the required axioms.

- *Conjugate transpose matrices*: Let $\mathbf{B} = (b_{ij})$ denote a matrix whose entries are complex numbers. Then the matrix $\mathbf{B}^* = (b_{ij}^*) = (\overline{b_{ji}})$ is called the *conjugate transpose* or *adjoint* of the matrix $\mathbf{B}$.

  We note the following facts:

1. Clearly, $\mathbf{B}^{**} = \mathbf{B}$, $(\mathbf{B}+\mathbf{C})^* = \mathbf{B}^* + \mathbf{C}^*$, and $(k\mathbf{B})^* = \bar{k}\mathbf{B}^*$.
2. $(\mathbf{AB})^* = \mathbf{B}^*\mathbf{A}^*$ and $(\mathbf{B}^n)^* = (\mathbf{B}^*)^n$.
3. $\langle \mathbf{Bx},\mathbf{y}\rangle = \langle \mathbf{x},\mathbf{B}^*\mathbf{y}\rangle$.
4. The determinants of $\mathbf{B}$ and $\mathbf{B}^*$ satisfy $\det(\mathbf{B}^*) = \overline{\det(\mathbf{B})}$.
5. If $\det(\mathbf{B}) \neq 0$, then $(\mathbf{B}^*)^{-1} = (\mathbf{B}^{-1})^*$.
6. The matrices $\mathbf{B}$ and $\mathbf{B}^*$ have complex conjugate eigenvalues.
7. If $\det(\mathbf{B}) \neq 0$, then both homogeneous systems of linear equations

$$\mathbf{Bx} = \mathbf{0} \quad \text{and} \quad \mathbf{B}^*\mathbf{y} = \mathbf{0}$$

   have only the trivial solution.
8. If $\det(\mathbf{B}) = 0$, then the ranks and nullities of the matrices $\mathbf{B}$ and $\mathbf{B}^*$ are equal. Consequently, both of the homogeneous systems above have the same number of linearly independent solutions.
9. If $\det(\mathbf{B}) = 0$, then the inhomogeneous system $\mathbf{Bx} = \mathbf{f}$ is solvable if and only if the vector $\mathbf{f}$ is orthogonal to all of the solutions of the homogeneous system $\mathbf{B}^*\mathbf{y} = \mathbf{0}$.

1.2 An Illustrative Example

In order to motivate the full theoretical treatment of the Fredholm integral equation of the second kind with a separable kernel in the next section, we examine in depth here an illustrative example.

Consider the equation

$$\phi(x) = f(x) + \lambda \int_0^1 \left(xt^2 + x^2t^4\right)\phi(t)\,\mathrm{d}t, \tag{1.1}$$

with real-valued kernel $K(x,t) = xt^2 + x^2t^4$. If we set

$$c_1 = \int_0^1 t^2\,\phi(t)\,\mathrm{d}t \quad \text{and} \quad c_2 = \int_0^1 t^4\,\phi(t)\,\mathrm{d}t,$$

then the integral equation can be reformulated as

$$\phi(x) = f(x) + \lambda\,c_1\,x + \lambda\,c_2\,x^2. \tag{1.2}$$

Observe that *any possible solution to the integral equation must assume this general form* but that the specific values of c_1 and c_2 are as yet undetermined.

After replacing x by t in Eq. (1.2), multiplying by t^2, and integrating the result from 0 to 1, we obtain

$$\int_0^1 t^2\,\phi(t)\,\mathrm{d}t = \int_0^1 t^2 f(t)\,\mathrm{d}t + \lambda\,c_1 \int_0^1 t^3\,\mathrm{d}t + \lambda\,c_2 \int_0^1 t^4\,\mathrm{d}t.$$

After replacing x by t in Eq. (1.2), multiplying by t^4, and integrating the result from 0 to 1, we obtain

$$\int_0^1 t^4\,\phi(t)\,\mathrm{d}t = \int_0^1 t^4 f(t)\,\mathrm{d}t + \lambda\,c_1 \int_0^1 t^5\,\mathrm{d}t + \lambda\,c_2 \int_0^1 t^6\,\mathrm{d}t.$$

If we set

$$f_1 = \int_0^1 t^2 f(t)\,\mathrm{d}t \quad \text{and} \quad f_2 = \int_0^1 t^4 f(t)\,\mathrm{d}t,$$

then these equations assume the form

$$\left(1 - \frac{\lambda}{4}\right) c_1 - \left(\frac{\lambda}{5}\right) c_2 = f_1 \tag{1.3}$$

and

$$\left(-\frac{\lambda}{6}\right) c_1 + \left(1 - \frac{\lambda}{7}\right) c_2 = f_2. \tag{1.4}$$

Algebraically, these equations constitute a linear system of two equations in the two unknowns, c_1 and c_2. Geometrically, they represent two lines in the c_1c_2-plane that may intersect, be parallel, or coincide, depending upon the values of λ, f_1, and f_2. All three of these possibilities may occur. This system can also be written in matrix form as

$$\begin{pmatrix} 1 - \frac{\lambda}{4} & -\frac{\lambda}{5} \\ -\frac{\lambda}{6} & 1 - \frac{\lambda}{7} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix},$$

or more elegantly as

$$(\mathbf{I} - \lambda\,\mathbf{A})\,\mathbf{c} = \mathbf{f}, \tag{1.5}$$

where we have set

$$\mathbf{A} = \begin{pmatrix} \frac{1}{4} & \frac{1}{5} \\ \frac{1}{6} & \frac{1}{7} \end{pmatrix}. \tag{1.6}$$

The analysis of this linear system depends upon the value of the *Fredholm determinant*

$$\det(\mathbf{I} - \lambda\,\mathbf{A}) = \frac{\lambda^2 - 165\lambda + 420}{420}. \tag{1.7}$$

There are exactly two values of λ for which this determinant vanishes. These two values are called *eigenvalues of the kernel* and are easily computed to be

$$\lambda_1 = \frac{165 - \sqrt{25545}}{2} = 2.585\ldots \quad \text{and} \quad \lambda_2 = \frac{165 + \sqrt{25545}}{2} = 162.414\ldots$$

Note: *A strong word of caution is in order here*! In Sect. 1.1, the eigenvalues of the matrix $\mathbf{A}$ were defined to be the values μ for which $\det(\mu\,\mathbf{I} - \mathbf{A}) = 0$. Thus, the eigenvalues of the kernel are precisely the *reciprocals* of the eigenvalues of the matrix, i.e., $\lambda = 1/\mu$.

We consider two cases:

Case I : $\det(\mathbf{I} - \lambda\,\mathbf{A}) \neq 0$.

In this case, Eqs. (1.3) and (1.4) represent two lines in the c_1c_2-plane that intersect in a single point. The unique solution $\mathbf{c}$ to the equivalent linear system (1.5) can be written in the form

$$\mathbf{c} = (\mathbf{I} - \lambda\,\mathbf{A})^{-1}\,\mathbf{f},$$

or more explicitly as

$$\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \frac{420}{\lambda^2 - 165\lambda + 420} \begin{pmatrix} 1 - \frac{\lambda}{7} & \frac{\lambda}{5} \\ \frac{\lambda}{6} & 1 - \frac{\lambda}{4} \end{pmatrix} \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}.$$

We consider three possibilities:

1. If $f(x) = 0$, then $f_1 = f_2 = 0$ so that $c_1 = c_2 = 0$. Since every solution to Eq. (1.1) has the form (1.2), it follows that the unique solution to Eq. (1.1) is the trivial solution $\phi(x) = 0$.
2. If $f(x) \neq 0$, but $f_1 = f_2 = 0$, then again $c_1 = c_2 = 0$. In this case, the unique solution to Eq. (1.1) is $\phi(x) = f(x)$.

3. If $f(x) \neq 0$ and either $f_1 \neq 0$ or $f_2 \neq 0$, then the definitions of f_1 and f_2 allow us to rewrite the expressions for c_1 and c_2 in the forms

$$c_1 = \frac{420}{\lambda^2 - 165\lambda + 420} \int_0^1 \left[\left(1 - \frac{\lambda}{7}\right) t^2 + \left(\frac{\lambda}{5}\right) t^4 \right] f(t)\,\mathrm{d}t$$

and

$$c_2 = \frac{420}{\lambda^2 - 165\lambda + 420} \int_0^1 \left[\left(\frac{\lambda}{6}\right) t^2 + \left(1 - \frac{\lambda}{4}\right) t^4 \right] f(t)\,\mathrm{d}t.$$

If we substitute these expressions for c_1 and c_2 into Eq. (1.2), then the unique solution to the inhomogeneous equation (1.1) can be represented in the form

$$\phi(x;\lambda) = f(x) + \lambda \int_0^1 R(x,t;\lambda) f(t)\,\mathrm{d}t, \tag{1.8}$$

where

$$R(x,t;\lambda) = \frac{420\left[\left(1 - \frac{\lambda}{7}\right) xt^2 + \left(\frac{\lambda}{5}\right) xt^4 + \left(\frac{\lambda}{6}\right) x^2 t^2 + \left(1 - \frac{\lambda}{4}\right) x^2 t^4\right]}{\lambda^2 - 165\lambda + 420}. \tag{1.9}$$

We write the solution as $\phi(x;\lambda)$ to indicate its dependence on the parameter λ. The function $R(x,t;\lambda)$ that appears in the integrand is called the *resolvent kernel* of the integral equation (1.1). The values of λ for which the resolvent exists are called *regular*.

For instance, if we choose $f(x) = x^3$, then the solution to the inhomogeneous integral equation (1.1) assumes the form

$$\phi(x;\lambda) = x^3 + \frac{\lambda}{\lambda^2 - 165\lambda + 420} \left[\frac{1}{2}(140 + \lambda)x + \frac{35}{24}(36 - \lambda)x^2 \right] \tag{1.10}$$

for every regular value of λ. It is worth noting here that

$$\lim_{\lambda \to \pm\infty} \phi(x;\lambda) = x^3 + \frac{1}{2}x - \frac{35}{24}x^2$$

but that the limits $\lim_{\lambda \to \lambda_i} \phi(x;\lambda)$ do not exist for $i = 1, 2$.

Case II: $\det(\mathbf{I} - \lambda\,\mathbf{A}) = 0$.

In this case, Eqs. (1.3) and (1.4) represent two parallel lines in the c_1c_2-plane. To see this, we multiply (1.3) by $(-\frac{\lambda}{6})$ and (1.4) by $(1 - \frac{\lambda}{4})$ to obtain

$$\left(-\frac{\lambda}{6}\right)\left(1 - \frac{\lambda}{4}\right) c_1 + \left(-\frac{\lambda}{6}\right)\left(-\frac{\lambda}{5}\right) c_2 = \left(-\frac{\lambda}{6}\right) f_1 \tag{1.11}$$

and

$$\left(1-\frac{\lambda}{4}\right)\left(-\frac{\lambda}{6}\right)c_1+\left(1-\frac{\lambda}{4}\right)\left(1-\frac{\lambda}{7}\right)c_2=\left(1-\frac{\lambda}{4}\right)f_2. \tag{1.12}$$

The coefficients of c_1 are equal. The coefficients of c_2 are also equal since

$$\det(\mathbf{I}-\lambda\,\mathbf{A})=\left(1-\frac{\lambda}{4}\right)\left(1-\frac{\lambda}{7}\right)-\left(-\frac{\lambda}{6}\right)\left(-\frac{\lambda}{5}\right)=0.$$

We now consider two possibilities:

1. If $(-\frac{\lambda}{6})f_1\neq(1-\frac{\lambda}{4})f_2$, then Eqs. (1.11) and (1.12) represent two distinct parallel lines in the c_1c_2-plane. In this situation, there are no solutions to the linear system and hence no solutions to the integral equation (1.1) either.
2. If $(-\frac{\lambda}{6})f_1=(1-\frac{\lambda}{4})f_2$, then Eqs. (1.11) and (1.12) represent two coincident lines in the c_1c_2-plane. In this situation, there are an infinite number of solutions to the linear system and hence an infinite number of solutions to the integral equation (1.1) as well for each eigenvalue.

   (a) If $f(x)=0$, then $f_1=f_2=0$, so that Eqs. (1.11) and (1.12) become

$$\left(1-\frac{\lambda}{4}\right)c_1=\left(\frac{\lambda}{5}\right)c_2\quad\text{or}\quad\left(\frac{\lambda}{6}\right)c_1=\left(1-\frac{\lambda}{7}\right)c_2.$$

   These two equations are identical. Recall that every solution to Eq. (1.1) must have the form (1.2). Since c_2 is a multiple of c_1, we can write the solution to the homogeneous integral equation in the form

$$\phi(x;\lambda)=\phi^{(e)}(x;\lambda)=c\left[\left(\frac{\lambda}{5}\right)x+\left(1-\frac{\lambda}{4}\right)x^2\right], \tag{1.13}$$

   where c is an arbitrary constant and λ may be either λ_1 or λ_2. Each of these solutions is called a *characteristic function* or an *eigenfunction of the kernel*. The superscript (e) indicates that the solution is an eigenfunction.

   (b) If $f(x)\neq0$ but $f_1=f_2=0$, then the analysis proceeds as in the previous case. We conclude that we can write the solution to the inhomogeneous integral equation in the form

$$\phi(x;\lambda)=f(x)+\phi^{(e)}(x;\lambda)=f(x)+c\left[\left(\frac{\lambda}{5}\right)x+\left(1-\frac{\lambda}{4}\right)x^2\right],$$

   where again c is an arbitrary constant and λ may be either λ_1 or λ_2.

   (c) If $f(x)\neq0$ with $f_1\neq0$ and $f_2\neq0$, then we have

$$c_2=\frac{\left(1-\frac{\lambda}{4}\right)c_1-f_1}{\left(\frac{\lambda}{5}\right)}$$

from Eq. (1.11). Substituting c_2 into the required form (1.2), we find that the solution to the inhomogeneous equation in this case is

$$\phi(x;\lambda) = f(x) + c\left[\left(\frac{\lambda}{5}\right)x + \left(1-\frac{\lambda}{4}\right)x^2\right] - 5f_1x^2,$$

where λ may be chosen to be either eigenvalue and $c = 5c_1$ is an arbitrary constant. Alternately, the solution may be expressed in the form

$$\phi(x;\lambda) = f(x) + c\,\phi^{(e)}(x;\lambda) + \phi^{(p)}(x;\lambda),$$

where $\phi^{(e)}(x;\lambda)$ is the eigenfunction from case (a) above and $\phi^{(p)}(x;\lambda)$ is some particular combination of x and x^2.

We have seen in Case II that the requirement that $(-\frac{\lambda}{6})f_1 = (1-\frac{\lambda}{4})f_2$ is a necessary and sufficient condition for the existence of a solution to the integral equation (1.1) when λ is an eigenvalue of the kernel. Consequently, it deserves further scrutiny. Recalling the definitions of f_1 and f_2, this requirement becomes

$$\int_0^1 \left[\left(\frac{\lambda}{6}\right)t^2 + \left(1-\frac{\lambda}{4}\right)t^4\right] f(t)\,\mathrm{d}t = 0. \tag{1.14}$$

Employing the vocabulary introduced in Sect. 1.1, this equation says that the free term $f(x)$ is *orthogonal* to the polynomial $(\frac{\lambda}{6})x^2 + (1-\frac{\lambda}{4})x^4$ on the interval $[0,1]$.

In order to explain the appearance of the polynomial in this condition, we now undertake an analysis of the homogeneous adjoint equation, which in our illustrative example assumes the form

$$\psi(x) = \lambda \int_0^1 \left(x^2t + x^4t^2\right)\psi(t)\,\mathrm{d}t. \tag{1.15}$$

If we set

$$d_1 = \int_0^1 t\,\psi(t)\,\mathrm{d}t \quad \text{and} \quad d_2 = \int_0^1 t^2\,\psi(t)\,\mathrm{d}t,$$

then the adjoint equation becomes

$$\psi(x) = \lambda\,d_1x^2 + \lambda\,d_2x^4. \tag{1.16}$$

Any possible solution to the adjoint equation must assume this form. The specific values of d_1 and d_2 must be determined.

Proceeding as before, we obtain the equations

$$\left(1-\frac{\lambda}{4}\right)d_1 - \left(\frac{\lambda}{6}\right)d_2 = 0 \quad \text{and} \quad \left(-\frac{\lambda}{5}\right)d_1 + \left(1-\frac{\lambda}{7}\right)d_2 = 0.$$

Equivalently, we have

$$\begin{pmatrix} 1-\frac{\lambda}{4} & -\frac{\lambda}{6} \\ -\frac{\lambda}{5} & 1-\frac{\lambda}{7} \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

or more compactly as

$$(\mathbf{I} - \lambda\, \mathbf{A}^{\mathrm{T}})\, \mathbf{d} = \mathbf{0},$$

where $\mathbf{A}^{\mathrm{T}}$ denotes the transpose of the matrix $\mathbf{A}$ defined above by (1.6). Since $\det \mathbf{B} = \det \mathbf{B}^{\mathrm{T}}$ for any real-valued square matrix $\mathbf{B}$, the eigenvalues of $\mathbf{A}$ and $\mathbf{A}^{\mathrm{T}}$ are the same. Thus, the eigenvalues of both corresponding kernels are the same as well. However, the eigenvectors of $\mathbf{A}$ and $\mathbf{A}^{\mathrm{T}}$ that correspond to the same eigenvalue may in general be different.

Recall that every solution to Eq. (1.15) must have the form (1.16). Since d_2 is a multiple of d_1, we can write the solution to the homogeneous adjoint equation in the form

$$\psi(x;\lambda) = \psi^{(e)}(x;\lambda) = d\left[\left(\frac{\lambda}{6}\right)x^2 + \left(1-\frac{\lambda}{4}\right)x^4\right],$$

where d is an arbitrary constant and λ may be either λ_1 or λ_2. Each of these solutions is an eigenfunction of the kernel appearing in the adjoint equation. Note especially that $\phi^{(e)}(x;\lambda) \neq \psi^{(e)}(x;\lambda)$.

The orthogonality condition (1.14) now becomes

$$\int_0^1 \psi^{(e)}(x;\lambda)\, f(t)\, \mathrm{d}t = 0.$$

Rephrased, this condition requires that the free term $f(x)$ be orthogonal to the eigenfunctions of the adjoint equation for a solution to exist to Eq. (1.1) when λ is an eigenvalue of the kernel.

It is instructive to highlight the essential features exhibited in an example because they usually enhance our understanding of the general theory to follow. Accordingly, we record the following observations:

- The resolvent kernel (1.9) is separable. It is worth noting that it can be written as the negative quotient of two determinants. Specifically, we have

$$R(x,t;\lambda) = -\frac{\det(\mathbf{D}(x,t;\lambda))}{\det(\mathbf{I}-\lambda\,\mathbf{A})}, \tag{1.17}$$

where

$$\mathbf{D}(x,t;\lambda) = \begin{pmatrix} 0 & x & x^2 \\ t^2 & 1-\frac{\lambda}{4} & -\frac{\lambda}{5} \\ t^4 & -\frac{\lambda}{6} & 1-\frac{\lambda}{7} \end{pmatrix}.$$

Note that the matrix $\mathbf{I} - \lambda\,\mathbf{A}$ is a submatrix of $\mathbf{D}(x,t;\lambda)$. Since $R(x,t;\lambda)$ is a quotient of two polynomials in λ, it can be developed in a power series in λ, whose radius of convergence is equal to the smallest eigenvalue λ_1.

- The value of the complex parameter λ plays a critical role in determining the existence and uniqueness of a solution to the integral equation and its adjoint equation. There are essentially two alternatives.

  If λ is not an eigenvalue of the kernel, then Eq. (1.1) has the unique solution (1.8) for any continuous free term $f(x)$. This fact illustrates *the First Fredholm Theorem*.

  If λ is an eigenvalue of the kernel, then the inhomogeneous equation (1.1) will have a solution if and only if the free term $f(x)$ is orthogonal to the eigenfunctions of the adjoint equation (1.15). This observation is known as *the Third Fredholm Theorem*.
- In the illustrative example, both of the kernels, $K(x,t) = xt^2 + x^2t^4$ and $K(t,x) = x^2t + x^4t^2$, had two eigenvalues and two corresponding eigenfunctions. The eigenvalues λ_i were the same, but the eigenfunctions, $\phi^{(e)}(x;\lambda_i)$ and $\psi^{(e)}(x;\lambda_i)$, were different. A separable kernel of the form

$$K(x,t) = \sum_{i=1}^{n} a_i(x)\, b_i(t)$$

  will have exactly n eigenvalues, counting multiplicities. If λ is any common eigenvalue of these kernels, then both the homogeneous Fredholm integral equation and its homogeneous adjoint equation will have the same finite number of linearly independent eigenfunctions corresponding to λ. This observation is known as *the Second Fredholm Theorem.*
- Separable kernels have a finite number of eigenvalues. However, arbitrary nonseparable kernels may have a countably infinite number of eigenvalues. The fact that these eigenvalues do not have a finite subsequential limit point is known as *the Fourth Fredholm Theorem*.

Section 1.2 Exercises

1. A function $\phi(x)$ is called a *solution* of an integral equation if the integral equation is transformed into an identity by substituting $\phi(x)$ into it. Show that $\phi(x) = 1$ is the solution to the Fredholm integral equation

$$\phi(x) = \mathrm{e}^x - x - \int_0^1 x(\mathrm{e}^{xt} - 1)\,\phi(t)\,\mathrm{d}t.$$

2. Solve each of the following inhomogeneous Fredholm integral equations of the second kind for all values of λ for which there is a solution:
   (a) $\phi(x) = x^3 + \lambda \int_0^1 xt\,\phi(t)\,\mathrm{d}t$
   (b) $\phi(x) = \sin(\pi x) + \frac{1}{2}\int_0^1 \phi(t)\,\mathrm{d}t$
   (c) $\phi(x) = x^2 + \lambda \int_0^1 \mathrm{e}^{x+t}\,\phi(t)\,\mathrm{d}t$
   (d) $\phi(x) = \cos x + \lambda \int_0^\pi \sin x\,\phi(t)\,\mathrm{d}t$
   (e) $\phi(x) = \sin x + \lambda \int_0^{2\pi} \sin(x/2)\,\phi(t)\,\mathrm{d}t$
   (f) $\phi(x) = \sin x + \lambda \int_0^{2\pi} \sin(t/2)\,\phi(t)\,\mathrm{d}t$
   (g) $\phi(x) = \sin x + \lambda \int_0^{2\pi} \sin(x/2)\sin(t/2)\,\phi(t)\,\mathrm{d}t$
   (h) $\phi(x) = \frac{5}{6}x + \frac{1}{2}\int_0^1 xt\,\phi(t)\,\mathrm{d}t$
   (i) $\phi(x) = x^4 + \lambda \int_0^1 x^2 t^2\,\phi(t)\,\mathrm{d}t$
   (j) $\phi(x) = 1 + \lambda \int_0^1 |x|\,\phi(t)\,\mathrm{d}t$
3. The simplest separable kernels have one of the forms $K(x,t) = a(x)$, $K(x,t) = b(t)$ or $K(x,t) = a(x)\,b(t)$. Given your experience with the previous problem, what can you say in general about the solutions to Fredholm integral equations with kernels of these forms?
4. Consider the integral equation

$$\phi(x) = 1 + \lambda \int_0^1 (x - t)\,\phi(t)\,\mathrm{d}t.$$

   (a) Show that the integral equation has a unique solution for every complex value of $\lambda \neq \pm 2\mathrm{i}\sqrt{3}$. What happens to this solution as $\lambda \to \infty$? As $\lambda \to \pm 2\mathrm{i}\sqrt{3}$?
   (b) Show that if $\lambda \neq \pm 2\mathrm{i}\sqrt{3}$, then the homogeneous equation

$$\phi(x) = \lambda \int_0^1 (x - t)\,\phi(t)\,\mathrm{d}t$$

   has only the trivial solution $\phi(x) \equiv 0$, but that if $\lambda = \pm 2\mathrm{i}\sqrt{3}$, then this equation has nontrivial solutions.
   (c) Show also that the inhomogeneous integral equation above has no solution if $\lambda = \pm 2\mathrm{i}\sqrt{3}$.

5. Solve each of the following inhomogeneous Fredholm integral equations of the second kind for all values of λ for which there is a solution:
   (a) $\phi(x) = \cos x + \lambda \int_0^\pi \mathrm{e}^x \cos t\,\phi(t)\,\mathrm{d}t$
   (b) $\phi(x) = 1 + x^2 + \lambda \int_0^1 (x + t)\,\phi(t)\,\mathrm{d}t$

(c) $\phi(x) = 1 + x^2 + \lambda \int_0^1 \frac{x}{1+t^2}\,\phi(t)\,\mathrm{d}t$

(d) $\phi(x) = 1 + \lambda \int_0^{\pi} x \cos^2 t\,\phi(t)\,\mathrm{d}t$

(e) $\phi(x) = x + \lambda \int_0^{\pi} (1 + \sin x \sin t)\,\phi(t)\,\mathrm{d}t$

(f) $\phi(x) = 1 + \lambda \int_{-\pi}^{\pi} \cos(x+t)\,\phi(t)\,\mathrm{d}t$

6. Invent an integral equation for which $\phi(x) = \mathrm{e}^x + \sin x$ is the solution.
7. Invent a kernel that is continuous on the square $Q(1,2)$ and has eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 2$. What are the eigenfunctions of your kernel?
8. Invent a kernel that is defined and continuous on the square $Q(a,b)$ and has the double eigenvalue $\lambda_1 = \lambda_2 = 1$. Does your invented kernel have two eigenfunctions corresponding to this eigenvalue or just one?
9. Show that the resolvent kernel (1.9) has the representation (1.17).
10. Our analysis of equation with the kernel $K(x,t) = xt^2 + x^2 t^4$ led us to the matrix $\mathbf{A}$ given by (1.6). If we substitute a generic kernel of the form $K(x,t) = a_1(x)\,b_1(t) + a_2(x)\,b_2(t)$, in Eq. (1.1), then our analysis will lead us to a generic form for $\mathbf{A} = (a_{ij})$. What are the entries in this generic form?
Answer: $a_{ij} = \int_a^b a_j(t)\,b_i(t)\,\mathrm{d}t$.
11. In the illustrative example, we set

$$a_1(x) = x, \qquad a_2(x) = x^2, \qquad b_1(t) = t^2, \quad \text{and} \quad b_2(t) = t^4$$

in the kernel $K(x,t) = a_1(x)\,b_1(t) + a_2(x)\,b_2(t)$. Suppose instead that we had set

$$\tilde{a}_1(x) = \frac{1}{2}x, \qquad \tilde{a}_2(x) = \frac{1}{4}x^2, \qquad \tilde{b}_1(t) = 2t^2, \quad \text{and} \quad \tilde{b}_2(t) = 4t^4.$$

Does this change of assignment affect the solutions to the integral equation in any way?
Answer: No. The matrices $\mathbf{A} = (a_{ij})$ and $\tilde{\mathbf{A}} = (\tilde{a}_{ij})$ are different, but the eigenvalues, the eigenfunctions, and the resolvent kernel are the same.

1.3 The Fredholm Theorems

In the previous section, we introduced algorithmic techniques for determining the solution to a particular Fredholm integral equation of the second kind in order to motivate the use of these methods in general.

In this section, we present the theoretical basis for a thorough discussion of the Fredholm integral equation of the second kind

$$\phi(x) = f(x) + \lambda \int_a^b K(x,t)\,\phi(t)\,\mathrm{d}t, \tag{1.18}$$

where $K(x,t)$ is an arbitrary *separable* kernel defined on $Q(a,b)$ of the form

$$K(x,t) = \sum_{i=1}^{n} a_i(x)\,\overline{b_i(t)}. \tag{1.19}$$

We assume that each function $a_i(x)$ and $b_i(t)$ is complex-valued and continuous on the interval $[a,b]$. We further assume without any loss of generality that the sets $\{a_i(x)\}$ and $\{b_i(t)\}$ are linearly independent on the interval $[a,b]$. The solution $\phi(x)$ is assumed to be Riemann integrable. The *free term* $f(x)$ is assumed to be complex-valued and continuous on the interval $[a,b]$. If $f(x) \equiv 0$ on the interval $[a,b]$, then the integral equation is called *homogeneous*; otherwise, it is called *inhomogeneous*. The complex constant $\lambda\ (\neq 0)$ is a parameter that should not be absorbed into the kernel. We will see that the form of the solution to the integral equation depends directly upon the value of λ.

Any complete treatment of Eq. (1.18) requires a thorough explanation of the significant role that the homogeneous *adjoint equation*[2]

$$\psi(x) = \overline{\lambda} \int_a^b \overline{K(t,x)}\,\psi(t)\,\mathrm{d}t \tag{1.20}$$

plays in the Fredholm theory. We will see that its role is a natural consequence of a well-known result in the theory of linear algebra.

In this section, we will state and prove the four Fredholm Theorems and the Fredholm Alternative Theorem for Fredholm integral equations of the second kind with a separable kernel. In the next chapter, we will prove them for equations with a *general* kernel, i.e., for an arbitrary nonseparable kernel. The proofs there will depend upon the results of this section.

The analytic procedure begins by substituting the separable kernel (1.19) into Eq. (1.18), thereby obtaining

$$\phi(x) = f(x) + \lambda \sum_{i=1}^{n} a_i(x) \left(\int_a^b \overline{b_i(t)}\,\phi(t)\,\mathrm{d}t \right).$$

If we set

$$c_i = \int_a^b \overline{b_i(t)}\,\phi(t)\,\mathrm{d}t,$$

for $i = 1,\ldots,n$, then the previous equation becomes

$$\phi(x) = f(x) + \lambda \sum_{i=1}^{n} c_i\,a_i(x). \tag{1.21}$$

[2]In the literature, the adjoint equation is also referred to as the *conjugate* equation, the *transposed* equation, or the *associated* equation.

It is important to observe here that *any possible solution to the integral equation must assume this general form.* Thus, the process of solving a Fredholm integral equation with a separable kernel is directly related to the process by which the constants c_i are determined.

[**Note**: Both the solution $\phi(x) = \phi(x;\lambda)$ and the constants $c_i = c_i(\lambda)$ depend upon the value of λ. However, we shall henceforth suppress this dependence for the sake of notational simplicity, unless the context requires it.]

If we replace x by t in Eq. (1.21), change the index of summation from i to j, multiply both sides of the equation by $\overline{b_i(t)}$, and then integrate the resulting expression from a to b, then we obtain

$$\int_a^b \overline{b_i(t)}\,\phi(t)\,\mathrm{d}t = \int_a^b \overline{b_i(t)}\,f(t)\,\mathrm{d}t + \lambda \sum_{j=1}^{n} c_j \left(\int_a^b a_j(t)\,\overline{b_i(t)}\,\mathrm{d}t \right)$$

for $i = 1,\ldots,n$. After introducing the simplifying notations

$$f_i = \int_a^b \overline{b_i(t)}\,f(t)\,\mathrm{d}t \quad \text{and} \quad a_{ij} = \int_a^b a_j(t)\,\overline{b_i(t)}\,\mathrm{d}t,$$

we obtain a system of n linear equations in the n unknowns c_i which can be written in the form

$$c_i = f_i + \lambda \sum_{j=1}^{n} a_{ij}\,c_j \tag{1.22}$$

for $i = 1,\ldots,n$. For a given value of λ, the solvability of this linear system correlates with the solvability of the integral equation. Accordingly, we proceed to analyze this linear system, which can be elegantly stated in matrix form as

$$(\mathbf{I} - \lambda\mathbf{A})\,\mathbf{c} = \mathbf{f}, \tag{1.23}$$

where $\mathbf{I}$ is the $n \times n$ identity matrix and $\mathbf{A} = (a_{ij})$ is the $n \times n$ matrix whose entries are defined above. The entries of the column vector $\mathbf{f} = (f_1,\ldots,f_n)^{\mathrm{T}}$ are defined above in terms of the free term $f(x)$. The required solution to the system is denoted by the column vector $\mathbf{c} = (c_1,\ldots,c_n)^{\mathrm{T}}$.

The determinant $D(\lambda) = \det(\mathbf{I} - \lambda\mathbf{A})$, which is known as the *Fredholm determinant*, is a polynomial of degree n in the variable λ. It always has n zeroes counting multiplicity by the Fundamental Theorem of Algebra. These zeroes are called *eigenvalues of the kernel.*

The analysis of the linear system (1.23) proceeds in remarkedly different ways, depending upon the value of the determinant $D(\lambda)$.

We consider two cases:

Case I: $D(\lambda) = \det(\mathbf{I} - \lambda\mathbf{A}) \neq 0$.

In this case, the linear system has the unique solution

$$\mathbf{c} = (\mathbf{I} - \lambda\mathbf{A})^{-1}\,\mathbf{f}. \tag{1.24}$$

We consider two possibilities:

1. If $\mathbf{f} = \mathbf{0}$, then $\mathbf{c} = \mathbf{0}$, i.e., $c_i = 0$ for all $i = 1, \ldots, n$. Since every solution to Eq. (1.18) has the form (1.21), it follows that the unique solution to the integral equation is $\phi(x) = f(x)$. (It is possible for $\mathbf{f} = \mathbf{0}$ with $f(x) \neq 0$ on $[a,b]$.) In particular, if $f(x) \equiv 0$ on $[a,b]$, i.e., if Eq. (1.18) is homogeneous, then the unique solution is the trivial solution $\phi(x) \equiv 0$.
2. If $\mathbf{f} \neq \mathbf{0}$, then $\mathbf{c} \neq \mathbf{0}$, i.e., $c_i \neq 0$ for at least one subscript i. Upon the substitution of the nonzero values c_i into the required form (1.21), we will have produced thereby the unique solution to the integral equation.

Although we have successfully solved the integral equation in this case, we have not yet produced a canonical representation for it. Toward this end, we express (1.24) in the form

$$\mathbf{c} = \frac{1}{D(\lambda)} \operatorname{adj}(\mathbf{I} - \lambda \mathbf{A})\, \mathbf{f},$$

where $\operatorname{adj}(\mathbf{I} - \lambda \mathbf{A}) = (D_{ji}(\lambda))$ is the transpose of the matrix of cofactors from $(\mathbf{I} - \lambda \mathbf{A})$. Consequently, each coefficient c_i has the representation

$$c_i = \frac{1}{D(\lambda)} \sum_{j=1}^{n} D_{ji}(\lambda)\, f_j.$$

Substituting this representation for c_i into the solution (1.21), we obtain

$$\phi(x) = f(x) + \lambda \sum_{i=1}^{n} \left(\frac{1}{D(\lambda)} \sum_{j=1}^{n} D_{ji}(\lambda)\, f_j \right) a_i(x).$$

After recalling the definition of f_j, the solution can be reformulated as

$$\phi(x) = f(x) + \lambda \int_a^b \left(\frac{1}{D(\lambda)} \sum_{i=1}^{n} \sum_{j=1}^{n} D_{ji}(\lambda)\, a_i(x) \overline{b_j(t)} \right) f(t)\, \mathrm{d}t$$

or as

$$\phi(x) = f(x) + \lambda \int_a^b R(x,t;\lambda)\, f(t)\, \mathrm{d}t, \tag{1.25}$$

where $R(x,t;\lambda)$ denotes the expression in parentheses in the integrand.

The function $R(x,t;\lambda)$ that appears in the integrand is called the *resolvent kernel* of the integral equation. It is clear that the resolvent kernel depends only on λ and the functions $a_i(x)$ and $\overline{b_i(t)}$ that appear in the definition of $K(x,t)$. The values of λ for which the resolvent kernel exists (the values for which $D(\lambda) \neq 0$) are called *regular*. From this representation for the solution, it is clear that solving the integral

equation for any regular value of λ and any arbitrary free term $f(x)$ is equivalent to producing the resolvent kernel. The denominator $D(\lambda)$ in the expression for $R(x,t;\lambda)$ is the Fredholm determinant. It can be shown that the numerator

$$\sum_{i=1}^{n}\sum_{j=1}^{n} D_{ji}(\lambda)\,a_i(x)\,\overline{b_j(t)} = -\det(\mathbf{D}(x,t;\lambda)),$$

where

$$\mathbf{D}(x,t;\lambda) = \begin{pmatrix} 0 & a_1(x) & a_2(x) & \cdots & a_n(x) \\ \overline{b_1(t)} & 1-\lambda a_{11} & -\lambda a_{12} & \cdots & -\lambda a_{1n} \\ \overline{b_2(t)} & -\lambda a_{21} & 1-\lambda a_{22} & \cdots & -\lambda a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \overline{b_n(t)} & -\lambda a_{n1} & -\lambda a_{n2} & \cdots & 1-\lambda a_{nn} \end{pmatrix}. \tag{1.26}$$

Note that the $n \times n$ lower right block of this $(n+1)\times(n+1)$ matrix is precisely $(\mathbf{I}-\lambda\mathbf{A})$. The determinant $D(x,t;\lambda) = \det(\mathbf{D}(x,t;\lambda))$ can be evaluated by computing the cofactor expansion along the first row and then computing the subsequent cofactor expansions along the first column. As a consequence, the resolvent kernel $R(x,t;\lambda)$ can be represented in the canonical form

$$R(x,t;\lambda) = -\frac{D(x,t;\lambda)}{D(\lambda)} \tag{1.27}$$

as the negative quotient of determinants, each of which is a polynomial in the variable λ. Also, $\deg(D(x,t;\lambda)) = n-1$ and $\deg(D(\lambda)) = n$.

The discussion above serves to establish the following result:

Theorem 1.3.1 (The First Fredholm Theorem). *Let λ be a complex parameter, $f(x)$ be a complex-valued continuous function defined on the interval $[a,b]$, and $K(x,t)$ be a complex-valued continuous separable kernel of the form (1.19) that is defined on the square $Q(a,b)$. Then the unique solution $\phi(x)$ to the Fredholm integral equation of the second kind,*

$$\phi(x) = f(x) + \lambda \int_a^b K(x,t)\,\phi(t)\,\mathrm{d}t,$$

has the representation

$$\phi(x) = f(x) + \lambda \int_a^b R(x,t;\lambda)\,f(t)\,\mathrm{d}t,$$

for every regular value of λ, where $R(x,t;\lambda)$ is the resolvent kernel defined by (1.27).

Case II: $D(\lambda) = \det(\mathbf{I} - \lambda \mathbf{A}) = 0$.

In this case, the linear system $(\mathbf{I} - \lambda \mathbf{A})\,\mathbf{c} = \mathbf{f}$ may or may not have solutions. As a consequence, the corresponding integral equation may or may not have solutions either.

We consider two possibilities:

1. If $\mathbf{f} = \mathbf{0}$, then the homogeneous linear system becomes

$$(\mathbf{I} - \lambda_k \mathbf{A})\,\mathbf{c} = \mathbf{0}$$

for each zero λ_k of the Fredholm determinant $D(\lambda) = 0$. This system has a certain number p_k of nontrivial linearly independent vector solutions $\mathbf{c}^{(j)}(\lambda_k)$, $j = 1, \ldots, p_k$, which take the form

$$\mathbf{c}^{(j)}(\lambda_k) = \begin{pmatrix} c_1^{(j)}(\lambda_k) \\ \vdots \\ c_n^{(j)}(\lambda_k) \end{pmatrix}.$$

The set

$$\mathbf{E}_{\lambda_k} = \{\text{span}\,\{\mathbf{c}^{(j)}(\lambda_k)\} \colon\ j = 1, \ldots, p_k\}$$

is a vector space, called the *eigenspace corresponding to the eigenvalue* λ_k, and $\dim(\mathbf{E}_{\lambda_k}) = p_k$. The number p_k is called the *geometric multiplicity of the eigenvalue* λ_k. If λ_k is an m_k-fold zero of the equation $D(\lambda) = 0$, then m_k is called the *algebraic multiplicity of the eigenvalue*. In any case, $1 \le p_k \le m_k \le n$.

The functions

$$\phi_j(x; \lambda_k) = f(x) + \lambda_k \sum_{i=1}^{n} c_i^{(j)}(\lambda_k)\, a_i(x),$$

obtained by substituting the coefficients $c_i^{(j)}(\lambda_k)$ into the Eq. (1.21), are nontrivial solutions of the integral equation. (It is possible for $\mathbf{f} = \mathbf{0}$ with $f(x) \neq 0$ on $[a, b]$.) If $f(x) \equiv 0$ on $[a, b]$, then each of the functions

$$\phi_j^{(e)}(x; \lambda_k) = \lambda_k \sum_{i=1}^{n} c_i^{(j)}(\lambda_k)\, a_i(x)$$

will be a nontrivial solution of the homogeneous integral equation

$$\phi(x) = \lambda_k \int_a^b K(x, t)\, \phi(t)\, \mathrm{d}t, \tag{1.28}$$

for $j = 1, \ldots, p_k$. These solutions are referred to as *eigenfunctions of the kernel corresponding to the eigenvalue* λ_k and are designated by the superscript (e). The set

$$\mathcal{E}_{\lambda_k} = \left\{ \text{span} \left\{ \phi_j^{(e)}(x;\lambda_k) \right\} : j = 1, \ldots, p_k \right\}$$

of all linear combinations of eigenfunctions corresponding to the eigenvalue λ_k is also a vector space, and $\dim(\mathcal{E}_{\lambda_k}) = p_k$. Since the homogeneous equation is linear, every element of $\mathcal{E}_{\lambda_k}$, i.e., every linear combination of the form

$$\phi^{(h)}(x;\lambda_k) = \sum_{j=1}^{p_k} \alpha_j \, \phi_j^{(e)}(x;\lambda_k),$$

where the α_j are arbitrary constants, is a solution to the homogeneous equation, as indicated by the superscript (h).

2. If $\mathbf{f} \neq \mathbf{0}$, then the linear system $(\mathbf{I} - \lambda \mathbf{A})\,\mathbf{c} = \mathbf{f}$ may or may not have solutions. Fortunately, the following criterion is available to assist us in our investigation.

Lemma 1.3.1. *Let* $\mathbf{B}$ *be a square matrix with complex entries, and let* $\mathbf{B}^*$ *be its conjugate transpose. If* $\det(\mathbf{B}) = 0$*, then the inhomogeneous system* $\mathbf{Bx} = \mathbf{f}$ *is solvable if and only if the vector* $\mathbf{f}$ *is orthogonal to all of the solutions of the homogeneous conjugate transpose system* $\mathbf{B}^*\mathbf{y} = \mathbf{0}$.

Proof. If the inhomogeneous system $\mathbf{Bx} = \mathbf{f}$ is solvable and $\mathbf{B}^*\mathbf{y} = \mathbf{0}$, then

$$\langle \mathbf{f}, \mathbf{y} \rangle = \langle \mathbf{Bx}, \mathbf{y} \rangle = \mathbf{y}^*\mathbf{Bx} = (\mathbf{B}^*\mathbf{y})^*\,\mathbf{x} = \langle \mathbf{x}, \mathbf{B}^*\mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{0} \rangle = 0.$$

Conversely, if $\mathbf{f}$ is orthogonal to all of the solutions of the homogeneous conjugate transpose system $\mathbf{B}^*\mathbf{y} = \mathbf{0}$ (or $\mathbf{y}^*\mathbf{B} = \mathbf{0}^*$), then $\mathbf{f}$ must belong to the range of the linear transformation that is represented by the matrix $\mathbf{B}$, that is, $\mathbf{f}$ must be a linear combination of the columns of $\mathbf{B}$. If the coefficients in this representation are x_i, then choose $\mathbf{x} = (x_1, \ldots, x_n)$. □

We will eventually apply this very useful lemma to conclude that the linear system $(\mathbf{I} - \lambda_k \mathbf{A})\,\mathbf{c} = \mathbf{f}$ is solvable if and only $\mathbf{f}$ is orthogonal to all of the solutions of the conjugate transpose system $(\mathbf{I} - \lambda_k \mathbf{A})^*\,\mathbf{d} = \mathbf{0}$ for each zero λ_k of the Fredholm determinant, but a thoughtful examination of each of these systems and their relations to Eq. (1.20) is required beforehand.

The system $(\mathbf{I} - \lambda_k \mathbf{A})^*\,\mathbf{d} = \mathbf{0}$ has a certain number q_k of nontrivial linearly independent vector solutions $\mathbf{d}^{(j)}(\lambda_k)$, $j = 1, \ldots, q_k$, which take the form

$$\mathbf{d}^{(j)}(\lambda_k) = \begin{pmatrix} d_1^{(j)}(\lambda_k) \\ \vdots \\ d_n^{(j)}(\lambda_k) \end{pmatrix}.$$

The set

$$\mathbf{F}_{\lambda_k} = \left\{ \text{span } \{\mathbf{d}^{(j)}(\lambda_k)\} : \; j = 1, \ldots, q_k \right\}$$

is the eigenspace corresponding to the eigenvalue λ_k, and $\dim(\mathbf{F}_{\lambda_k}) = q_k$, where q_k is the geometric multiplicity of the eigenvalue λ_k. In any case, we always have $1 \leq q_k \leq m_k \leq n$.

Since the matrices $(\mathbf{I} - \lambda_k \mathbf{A})$ and $(\mathbf{I} - \lambda_k \mathbf{A})^*$ have the same ranks and nullities, it must necessarily be the case that $p_k = q_k$, i.e., that the eigenbases for $\mathbf{E}_{\lambda_k}$ and $\mathbf{F}_{\lambda_k}$ have the same cardinality.

The matrix equation

$$(\mathbf{I} - \lambda_k \mathbf{A})^* \, \mathbf{d} = \left(\mathbf{I} - \overline{\lambda_k} \mathbf{A}^* \right) \mathbf{d} = \left(\mathbf{I} - \overline{\lambda_k} \, \overline{\mathbf{A}}^{\mathrm{T}} \right) \mathbf{d} = \mathbf{0}$$

can be expanded as a system of n linear equations in n unknowns. Specifically, if $\mathbf{d}^{(j)}(\lambda_k)$ is any one of the p_k solutions to this system, then the expanded form of this system becomes

$$d_m^{(j)}(\lambda_k) - \overline{\lambda_k} \sum_{i=1}^{n} \overline{a_{im}} \, d_i^{(j)}(\lambda_k) = 0 \tag{1.29}$$

for $m = 1, \ldots, n$. The similarities between the expanded form of this system and the expanded form (1.22) are remarkable: the parameter λ there is replaced with $\overline{\lambda_k}$ here; the sum of the terms $a_{ij} c_j(\lambda)$ there over the second subscript is replaced with the sum of the terms $\overline{a_{im}} \, d_i^{(j)}(\lambda_k)$ here over the first subscript. Since Eq. (1.22) was derived from the integral equation (1.18), our comparison suggests that we consider another integral equation that could serve as a suitable companion to it.

With the kernel

$$\overline{K(t,x)} = \sum_{i=1}^{n} \overline{a_i(t)} \, b_i(x), \tag{1.30}$$

the homogeneous integral equation

$$\psi(x) = \overline{\lambda} \int_a^b \overline{K(t,x)} \, \psi(t) \, \mathrm{d}t, \tag{1.31}$$

called the *adjoint equation*, has precisely the desired characteristics. Indeed, after substituting the kernel $\overline{K(t,x)}$ into the adjoint equation, we obtain

$$\psi(x) = \overline{\lambda} \sum_{i=1}^{n} d_i \, b_i(x),$$

where

$$d_i = \int_a^b \overline{a_i(t)} \, \psi(t) \, \mathrm{d}t.$$

If we replace x by t, multiply both sides of the equation by $\overline{a_m(t)}$, and then integrate the resulting expression from a to b, we obtain the linear system

$$d_m - \overline{\lambda} \sum_{i=1}^{n} \overline{a_{im}}\, d_i = 0,$$

for $m = 1, \ldots, n$, where

$$\overline{a_{im}} = \int_a^b \overline{a_m(t)}\, b_i(t)\, \mathrm{d}t.$$

With λ replacing λ_k and $\mathbf{d} = (d_1, \ldots, d_n)^{\mathrm{T}}$ replacing $\mathbf{d}^{(j)}(\lambda_k)$, this linear system correlates precisely with the linear system (1.29). Furthermore, it follows from this correlation that the solutions to the homogeneous adjoint equation

$$\psi(x) = \overline{\lambda_k} \int_a^b \overline{K(t,x)}\, \psi(t)\, \mathrm{d}t \tag{1.32}$$

have the form

$$\psi_j^{(e)}(x; \overline{\lambda_k}) = \overline{\lambda_k} \sum_{i=1}^{n} d_i^{(j)}(\lambda_k)\, b_i(x)$$

for $j = 1, \ldots, p_k$. Since the homogeneous adjoint equation is linear, the set

$$\mathcal{F}_{\lambda_k} = \{\mathrm{span}\{\psi_j^{(e)}(x; \overline{\lambda_k})\} \colon\ j = 1, \ldots, p_k\}$$

of all linear combinations of eigenfunctions corresponding to the eigenvalue λ_k is also a vector space, and $\dim(\mathcal{F}_{\lambda_k}) = p_k$. Every element of $\mathcal{F}_{\lambda_k}$, i.e., every linear combination of the form

$$\psi^{(h)}(x; \overline{\lambda_k}) = \sum_{j=1}^{p_k} \beta_j\, \psi_j^{(e)}(x; \overline{\lambda_k}),$$

where the β_j are arbitrary constants, is a solution to the homogeneous equation, as indicated by the superscript (h).

The discussion above serves to establish the following result:

Theorem 1.3.2 (The Second Fredholm Theorem). *Let λ be a complex parameter and let $K(x,t)$ be a complex-valued continuous separable kernel defined on the square $Q(a,b)$. Then λ is an eigenvalue of the separable kernel $K(x,t)$ if and only if $\overline{\lambda}$ is an eigenvalue of the adjoint kernel $\overline{K(t,x)}$. The homogeneous integral equation*

$$\phi(x) = \lambda \int_a^b K(x,t)\, \phi(t)\, \mathrm{d}t$$

and the homogeneous adjoint integral equation

$$\psi(x) = \overline{\lambda} \int_a^b \overline{K(t,x)}\, \psi(t)\, \mathrm{d}t$$

have the same finite number of linearly independent eigenfunctions.

Proof. By definition, the zeroes of the equation

$$D^*(\lambda) = \det\left[(\mathbf{I} - \lambda \mathbf{A})^*\right] = \det\left(\mathbf{I} - \overline{\lambda}\mathbf{A}^*\right) = 0$$

are the eigenvalues of the kernel $\overline{K(t,x)}$. It also follows from this same equation that λ is an eigenvalue of the matrix $\mathbf{A}$ if and only if $\overline{\lambda}$ is an eigenvalue of the matrix $\mathbf{A}^*$. Hence, the first assertion follows. The second assertion follows since both equations have exactly p_k linearly independent eigenfunctions. □

We now turn to the linear system $(\mathbf{I} - \lambda_k \mathbf{A})\,\mathbf{c} = \mathbf{f}$. We assume that this linear system is solvable and then note the structure of all possible solutions. If $\mathbf{c}(\lambda_k) = (c_1(\lambda_k), \ldots, c_n(\lambda_k))^{\mathrm{T}}$ is the most general solution to the linear system and $\mathbf{c}^{(p)}(\lambda_k) = (c_1^{(p)}(\lambda_k), \ldots, c_n^{(p)}(\lambda_k))^{\mathrm{T}}$ denotes one solution in particular, then a simple subtraction shows that

$$(\mathbf{I} - \lambda_k \mathbf{A})\left(\mathbf{c}(\lambda_k) - \mathbf{c}^{(p)}(\lambda_k)\right) = \mathbf{0},$$

from which it follows that $\mathbf{c}(\lambda_k) - \mathbf{c}^{(p)}(\lambda_k)$ belongs to the eigenspace $\mathbf{E}_{\lambda_k}$. Hence, the most general solution to the linear system has the form

$$\mathbf{c}(\lambda_k) = \mathbf{c}^{(p)}(\lambda_k) + \sum_{j=1}^{p_k} \alpha_j\, \mathbf{c}^{(j)}(\lambda_k),$$

where the α_j are arbitrary constants. If we substitute this representation for the coefficients into the required form for the solution to the inhomogeneous integral equation, then we obtain

$$\phi(x) = f(x) + \lambda_k\, \phi^{(p)}(x;\lambda_k) + \beta\, \phi^{(h)}(x;\lambda_k), \tag{1.33}$$

where β is an arbitrary constant,

$$\phi^{(p)}(x;\lambda_k) = \sum_{i=1}^{n} c_i^{(p)}(\lambda_k)\, a_i(x) \quad \text{and} \quad \phi^{(h)}(x;\lambda_k) = \sum_{j=1}^{p_k} \alpha_j\, \phi_j^{(e)}(x;\lambda_k).$$

We can now apply the lemma stated earlier in this section to the problem of finding the solution to the inhomogeneous Fredholm integral equation when λ is an eigenvalue of the separable kernel $K(x,t)$. We interpret the lemma to say that the linear system $(\mathbf{I} - \lambda_k \mathbf{A})\,\mathbf{c} = \mathbf{f}$ is solvable if and only $\mathbf{f}$ is orthogonal to all of the solutions of the conjugate transpose system $(\mathbf{I} - \lambda_k \mathbf{A})^*\,\mathbf{d} = \mathbf{0}$ for each zero λ_k of the Fredholm determinant.

If we recall the definition

$$f_i = \int_a^b f(t)\,\overline{b_i(t)}\,\mathrm{d}t,$$

then, for each $j = 1,\ldots,p_k$, the orthogonality conditions imposed by the lemma become

$$\left\langle \mathbf{f}, \mathbf{d}^{(j)}(\lambda_k) \right\rangle = \sum_{i=1}^{n} f_i\,\overline{d_i^{(j)}(\lambda_k)} = \int_a^b f(t)\,\overline{\psi_j^{(e)}\left(t;\overline{\lambda_k}\right)}\,\mathrm{d}t = 0. \tag{1.34}$$

In other words, this condition requires that the free term $f(x)$ be orthogonal to all of the eigenfunctions of the kernel $\overline{K(t,x)}$. If this condition is fulfilled, then the inhomogeneous Fredholm integral equation will have solutions of the form (1.33). If, on the other hand, this condition is not fulfilled for any single value of j, then the inhomogeneous Fredholm integral equation will not have a solution.

The discussion above serves to establish the following result:

Theorem 1.3.3 (The Third Fredholm Theorem). *Let λ be a complex parameter, $f(x)$ be a complex-valued continuous function defined on the interval $[a,b]$, and $K(x,t)$ be a complex-valued continuous separable kernel defined on the square $Q(a,b)$. If λ is an eigenvalue of the kernel $K(x,t)$, then the inhomogeneous Fredholm integral equation of the second kind*

$$\phi(x) = f(x) + \lambda \int_a^b K(x,t)\,\phi(t)\,\mathrm{d}t$$

will have a solution if and only if the free term $f(x)$ is orthogonal to all of the eigenfunctions of the homogeneous adjoint equation

$$\psi(x) = \overline{\lambda} \int_a^b \overline{K(t,x)}\,\psi(t)\,\mathrm{d}t.$$

For general kernels, there is a Fourth Fredholm Theorem which states that the set of all eigenvalues of a kernel is at most countable and that this set of eigenvalues does not have a finite limit point. However, since separable kernels have only a finite number of eigenvalues, this theorem is trivially true here.

The essential features of the first three Fredholm theorems are combined into one comprehensive statement.

Theorem 1.3.4 (The Fredholm Alternative Theorem). *Let λ be a given complex parameter, $f(x)$ be a complex-valued continuous function defined on the interval $[a,b]$, and $K(x,t)$ be a complex-valued continuous separable kernel that is defined on the square $Q(a,b)$. Then*

either *the inhomogeneous Fredholm integral equation of the second kind*

$$\phi(x) = f(x) + \lambda \int_a^b K(x,t)\,\phi(t)\,\mathrm{d}t$$

possesses a unique solution (*if* λ *is not an eigenvalue of the kernel* $K(x,t)$)

or *the homogeneous equation*

$$\phi(x) = \lambda \int_a^b K(x,t)\,\phi(t)\,\mathrm{d}t$$

possesses nontrivial solutions (*if* λ *is an eigenvalue of the kernel* $K(x,t)$).
Furthermore, if λ *is an eigenvalue of the kernel* $K(x,t)$*, then the inhomogeneous equation possesses nontrivial solutions if and only if*

$$\int_a^b f(t)\,\overline{\psi(t)}\,\mathrm{d}t = 0$$

for every solution $\psi(x)$ *of the homogeneous adjoint equation*

$$\psi(x) = \overline{\lambda} \int_a^b \overline{K(t,x)}\,\psi(t)\,\mathrm{d}t.$$

A Concise Guide to Computation

- *Determine the eigenvalues of the kernel* $K(x,t)$:
  1. Compute the n^2 entries a_{ij} of the matrix $\mathbf{A}$ by using the formula

$$a_{ij} = \int_a^b a_j(t)\,\overline{b_i(t)}\,\mathrm{d}t.$$

  2. Form the matrix $(\mathbf{I} - \lambda\,\mathbf{A})$.
  3. Evaluate the Fredholm determinant, $D(\lambda) = \det(\mathbf{I} - \lambda\,\mathbf{A})$ which will be a polynomial of degree n in λ.
  4. The eigenvalues of the kernel are the zeroes of the equation $D(\lambda) = 0$.
- *Determine the resolvent kernel* $R(x,t;\lambda)$:
  1. Use the matrix $(\mathbf{I} - \lambda\,\mathbf{A})$ constructed above to construct the larger matrix

$$\mathbf{D}(x,t;\lambda) = \begin{pmatrix} 0 & a_1(x) & a_2(x) & \cdots & a_n(x) \\ \overline{b_1(t)} & 1-\lambda a_{11} & -\lambda a_{12} & \cdots & -\lambda a_{1n} \\ \overline{b_2(t)} & -\lambda a_{21} & 1-\lambda a_{22} & \cdots & -\lambda a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \overline{b_n(t)} & -\lambda a_{n1} & -\lambda a_{n2} & \cdots & 1-\lambda a_{nn} \end{pmatrix}.$$

Note that the $n \times n$ lower right block of this $(n+1)\times(n+1)$ matrix is precisely $(\mathbf{I} - \lambda \mathbf{A})$.

2. Evaluate the determinant $D(x,t;\lambda) = \det(\mathbf{D}(x,t;\lambda))$ by computing the cofactor expansion along the first row and then computing the subsequent cofactor expansions along the first column, thereby obtaining a linear combination of the functions $a_i(x)\overline{b_j(t)}$.
3. Represent the resolvent kernel $R(x,t;\lambda)$ in the canonical form

$$R(x,t;\lambda) = -\frac{D(x,t;\lambda)}{D(\lambda)}.$$

- *If λ is regular, formulate the solution to the inhomogeneous equation*:

1. If λ is a regular value, i.e., if λ is not an eigenvalue of the kernel, then substitute the resolvent kernel computed above into the equation

$$\phi(x) = f(x) + \lambda \int_a^b R(x,t;\lambda)\,\phi(t)\,\mathrm{d}t.$$

2. Evaluate the resulting integral to obtain the unique solution.

- *Determine all of the eigenfunctions of the kernel $K(x,t)$*:

1. For each eigenvalue λ_k of the kernel $K(x,t)$, solve the matrix equation $(\mathbf{I} - \lambda\,\mathbf{A})\,\mathbf{c} = \mathbf{0}$ for all of its solutions $\mathbf{c}^{(j)}(\lambda_k)$, $j = 1,\ldots,p_k$.
2. Use these solutions $\mathbf{c}^{(j)}(\lambda_k) = (c_1^{(j)}(\lambda_k),\ldots,c_n^{(j)}(\lambda_k))^{\mathrm{T}}$ to formulate all of the eigenfunctions

$$\phi_j^{(e)}(x;\lambda_k) = \lambda_k \sum_{i=1}^{n} c_i^{(j)}(\lambda_k)\,a_i(x)$$

for $j = 1,\ldots,p_k$. The span of these eigenfunctions constitutes all possible solutions to the homogeneous integral equation.

- *Determine all of the eigenfunctions of the kernel $\overline{K(t,x)}$*:

1. For each eigenvalue λ_k of the kernel $K(x,t)$, solve the matrix equation $(\mathbf{I} - \lambda\,\mathbf{A})^*\,\mathbf{d} = \mathbf{0}$ for all of its solutions $\mathbf{d}^{(j)}(\lambda_k)$, $j = 1,\ldots,p_k$.

2. Use these solutions $\mathbf{d}^{(j)}(\lambda_k) = (d_1^{(j)}(\lambda_k), \ldots, d_n^{(j)}(\lambda_k))^{\mathrm{T}}$ to formulate all of the eigenfunctions

$$\psi_j^{(e)}(x;\overline{\lambda_k}) = \overline{\lambda_k} \sum_{i=1}^{n} d_i^{(j)}(\lambda_k)\, b_i(x)$$

for $j = 1, \ldots, p_k$. The span of these eigenfunctions constitutes all possible solutions to the homogeneous adjoint equation.

- *Determine whether the free term is orthogonal to all of the solutions to the homogeneous adjoint equation*:

1. Evaluate the integrals

$$\int_a^b f(t)\, \overline{\psi_j^{(e)}(t;\overline{\lambda_k})}\, \mathrm{d}t.$$

2. If any of these integrals does not vanish, then stop here. If these integrals vanish for all j and k, then proceed to the next step.

- *If λ is an eigenvalue of the kernel, then formulate all of the solutions to the inhomogeneous integral equation*:

1. Determine any one particular solution $\mathbf{c}^{(p)}(\lambda_k)$ to the matrix system $(\mathbf{I} - \lambda\, \mathbf{A})\, \mathbf{c} = \mathbf{f}$ for eigenvalue λ_k.
2. Use the solution $\mathbf{c}^{(p)}(\lambda_k)$ of the matrix equation to form the particular solution $\phi^{(p)}(x;\lambda_k)$.
3. Formulate the solution to the inhomogeneous integral equation by adding $f(x)$, an arbitrary multiple β of the previously determined homogeneous solution, and the particular solution to obtain

$$\phi(x;\lambda_k) = f(x) + \beta\, \phi^{(h)}(x;\lambda_k) + \lambda_k\, \phi^{(p)}(x;\lambda_k).$$

Section 1.3 Exercises

1. Determine the unique solution of the integral equation

$$\phi(x) = \mathrm{e}^{2x} + \int_0^1 \left(x\mathrm{e}^t + \mathrm{e}^x t\right) \phi(t)\, \mathrm{d}t.$$

2. Determine the unique solution of the integral equation

$$\phi(x) = x^4 + \tfrac{1}{10} \int_0^1 \left(1 + xt + x^2 t^2\right) \phi(t)\, \mathrm{d}t.$$

3. Solve the Fredholm integral equation

$$\phi(x) = x^6 + \lambda \int_0^1 \left[x^5(1-t)^4 + x^4(1-t)^3\right] \phi(t)\,\mathrm{d}t$$

for all regular values of λ.

Hint: It will be helpful to know that

$$\int_0^1 t^{m-1}\,(1-t)^{n-1}\,\mathrm{d}t = \frac{\Gamma(m)\,\Gamma(n)}{\Gamma(m+n)}$$

and that $\Gamma(n) = (n-1)!$, where $\Gamma(n)$ is the Euler gamma function.

Solve the homogeneous equation

$$\phi(x) = \lambda_k \int_0^1 [x^5(1-t)^4 + x^4(1-t)^3]\,\phi(t)\,\mathrm{d}t$$

for every eigenvalue of the kernel.

4. Solve the Fredholm integral equation

$$\phi(x) = \mathrm{e}^x + \lambda \int_0^1 (x-t)^2\,\phi(t)\,\mathrm{d}t$$

for all regular values of λ. Solve the corresponding homogeneous equation

$$\phi(x) = \lambda \int_0^1 (x-t)^2\,\phi(t)\,\mathrm{d}t$$

for every eigenvalue of the kernel.

5. Solve the Fredholm integral equation

$$\phi(x) = x + \lambda \int_{-\pi}^{+\pi} \cos^2(x-t)\,\phi(t)\,\mathrm{d}t$$

for all regular values of λ. Solve the corresponding homogeneous equation

$$\phi(x) = \lambda_k \int_{-\pi}^{+\pi} \cos^2(x-t)\,\phi(t)\,\mathrm{d}t$$

for every eigenvalue of the kernel.

6. Let n be a positive integer and consider the homogeneous equation

$$\phi(x) = \lambda \int_{-\pi}^{+\pi} \cos^n(x-t)\,\phi(t)\,\mathrm{d}t.$$

Show that all of the eigenvalues of the kernel are of the form

$$\lambda_r = \frac{2^{n-1}}{\pi \binom{n}{r}}$$

and that all eigenfunctions of the kernel are of the form

$$1, \qquad \cos((n-2r)x) \quad \text{and} \quad \sin((n-2r)x),$$

where r assumes all nonnegative integer values not exceeding $n/2$. There will be $n+1$ such eigenfunctions. Show also that

$$\sum_{r=0}^{n} \frac{1}{\lambda_r} = 2\pi$$

for each positive integer n.

Hint: The kernel may be rewritten in a more convenient form since

$$\begin{aligned}\cos^n\theta &= \left(\frac{\mathrm{e}^{\mathrm{i}\theta} + \mathrm{e}^{-\mathrm{i}\theta}}{2}\right)^n \\ &= \frac{1}{2^n} \sum_{r=0}^{n} \binom{n}{r} \mathrm{e}^{\mathrm{i}(n-r)\theta}\,\mathrm{e}^{-\mathrm{i}r\theta} \\ &= \sum_{r=0}^{n} \frac{1}{2^n} \binom{n}{r} \cos(n-2r)\theta.\end{aligned}$$

This sum may be simplified according to whether n is even or odd.

7. It was shown in this section that the resolvent kernel $R(x,t;\lambda)$ can be written as the negative quotient of two determinants that, when expanded, are polynomials in the variable λ. For example, if the kernel has the form $K(x,t) = a_1(x)\overline{b_1(t)} + a_2(x)\overline{b_2(t)}$, then

$$R(x,t;\lambda) = -\frac{r_0(x,t) + r_1(x,t)\,\lambda}{(1-\lambda/\lambda_1)(1-\lambda/\lambda_2)}.$$

Determine $r_0(x,t)$ and $r_1(x,t)$. If $\lambda_1 \neq \lambda_2$, then the method of partial fractions can be used to write the resolvent kernel in the form

$$R(x,t;\lambda) = \frac{s_1(x,t)}{1-\lambda/\lambda_1} + \frac{s_2(x,t)}{1-\lambda/\lambda_2}.$$

Determine $s_1(x,t)$ and $s_2(x,t)$.

8. For $-1 \le x \le +1$, consider the polynomials

$$p_1(x) = x, \quad p_2(x) = 3x^2 - 1, \quad \text{and} \quad p_3(x) = 5x^3 - 3x.$$

Define the kernel

$$K(x,t) = \sum_{i=1}^{3} p_i(x)\, p_i(t)$$

on the square $Q(-1,+1)$.

(a) Set $a_{ij} = \int_{-1}^{+1} p_j(t)\, p_i(t)\, \mathrm{d}t$. Show that $a_{11} = \frac{2}{3}$, $a_{22} = \frac{8}{5}$, $a_{33} = \frac{8}{7}$, and that $a_{ij} = 0$ if $i \neq j$.

(b) For $i = 1,2,3$, determine the eigenvalues λ_i of the kernel.

(c) For $i = 1,2,3$, show that the eigenfunctions of the kernel are given by $\phi^{(e)}(x;\lambda_i) = c_i\, p_i(x)$ so that $\langle \phi^{(e)}(x;\lambda_i), \phi^{(e)}(x;\lambda_j) \rangle = 0$, if $i \neq j$.

(d) If $\lambda \neq \lambda_i$, show that the solution to the inhomogeneous Fredholm integral equation

$$\phi(x) = f(x) + \lambda \int_{-1}^{+1} K(x,t)\, \phi(t)\, \mathrm{d}t$$

is given by

$$\phi(x) = f(x) - \lambda \int_{-1}^{+1} \left(\sum_{i=1}^{3} \frac{p_i(x)\, p_i(t)}{1 - \lambda/\lambda_i} \right) f(t)\, \mathrm{d}t.$$

(e) Let $p_4(x) = 35x^4 - 30x^2 + 3$. Show that $p_4(x)$ is orthogonal to the eigenfunctions of the kernel, i.e., $p_4(x)$ is orthogonal to each $p_i(x)$, and then solve the inhomogeneous equation

$$\phi(x) = p_4(x) + \lambda_i \int_{-1}^{+1} K(x,t)\, \phi(t)\, \mathrm{d}t$$

for each eigenvalue λ_i.

9. Let n be a positive integer. For $m = 1,2,3$, and 4, solve each of the homogeneous Fredholm integral equations

$$\phi(x) = \lambda \int_{-\pi}^{+\pi} T_m(x,t)\, \phi(t)\, \mathrm{d}t,$$

where

$$T_1(x,t) = \cos(n(x-t)),$$
$$T_2(x,t) = \cos(n(x+t)),$$
$$T_3(x,t) = \sin(n(x-t)),$$

and

$$T_4(x,t) = \sin(n(x+t)).$$

Chapter 2
Fredholm Integral Equations of the Second Kind (General Kernel)

In Chap. 1, we conducted a thorough examination of the Fredholm integral equation of the second kind for an arbitrary complex parameter λ, assuming that the free term $f(x)$ is complex-valued and continuous on the interval $[a,b]$ and that the kernel $K(x,t)$ is complex-valued, continuous, and separable on the square $Q(a,b) = \{(x,t)\colon [a,b] \times [a,b]\}$. We stated the four Fredholm theorems and the Fredholm Alternative Theorem which provide for the construction of all possible solutions to the equation under these assumptions.

A question naturally arises: What, if anything, can be proven if $K(x,t)$ is a *general* kernel, i.e., an arbitrary kernel that is only assumed to be complex-valued and continuous? In this chapter, we will answer this question completely by proving that all of the Fredholm theorems continue to hold in this eventuality.

In Sect. 2.1, we present several tools of the trade which are indispensible for the comprehension of the material in this chapter.

In Sects. 2.2 and 2.3, we use these tools to show that the Fredholm integral equation of the second kind with a general kernel has a unique solution if the product of the parameter λ and the "size" of the kernel is small.

In Sect. 2.4, we prove the validity of the Fredholm theorems for unrestricted λ and a general kernel.

In Sect. 2.5, we show how to construct the resolvent kernel that appears in the solution to the integral equation recursively.

In Sect. 2.6, we introduce numerical methods for producing an approximation to the solution of a Fredholm integral equation. These methods are necessary due to the inherent computational difficulties in constructing the resolvent kernel.

2.1 Tools of the Trade

In this chapter, the reader should be familiar with the following topics:

- *Norms of continuous functions*: Let $\mathcal{C}[a,b]$ denote the vector space that consists of all complex-valued continuous functions on the interval $[a,b]$. A *norm* on $\mathcal{C}[a,b]$

S.M. Zemyan, *The Classical Theory of Integral Equations: A Concise Treatment*,
DOI 10.1007/978-0-8176-8349-8_2, © Springer Science+Business Media, LLC 2012

is a real-valued function ($f \in \mathcal{C}[a,b] \to \|f\| \in \mathbb{R}$) with the following properties:

1. $\|f\| \geq 0$ for every $f(x) \in \mathcal{C}[a,b]$
2. $\|f\| = 0$ if and only if $f(x) \equiv 0$ on $[a,b]$
3. $\|kf\| = |k|\,\|f\|$
4. $\|f+g\| \leq \|f\| + \|g\|$

In particular, we will have a strong interest in the *p-norm* $(0 < p < +\infty)$ defined by

$$\|f\|_p = \left(\int_a^b |f(x)|^p \, \mathrm{d}x \right)^{1/p}$$

in the special cases $p = 1$ and $p = 2$. Of critical importance is the *Cauchy–Schwarz inequality* which states that

$$|\langle f, g \rangle| \leq \|f\|_2 \, \|g\|_2$$

or, more explicitly, that

$$\left| \int_a^b f(x)\overline{g(x)} \, \mathrm{d}x \right| \leq \left(\int_a^b |f(x)|^2 \, \mathrm{d}x \right)^{1/2} \left(\int_a^b |g(x)|^2 \, \mathrm{d}x \right)^{1/2}.$$

Another norm of interest is the *supremum norm* defined by

$$\|f\|_\infty = \sup\{\, |f(x)| \colon x \in [a,b] \,\}.$$

In an entirely similar way, a norm can be defined on the set of complex-valued continuous functions defined on the square $Q(a,b)$. Kernels, in particular, can be normed. The norms

$$\|K\|_2 = \left(\int_a^b \int_a^b |K(x,t)|^2 \, \mathrm{d}x \, \mathrm{d}t \right)^{1/2}$$

and

$$\|K\|_\infty = \sup\{\, |K(x,t)| \colon (x,t) \in Q(a,b) \,\}$$

will be of special interest in this chapter.

- *Uniform convergence of an infinite sequence of functions*: An infinite sequence $\{f_n(x)\}$ of functions *converges uniformly* on the interval $[a,b]$ to a function $f(x)$ if, for every $\varepsilon > 0$, there exists an integer $N = N(\varepsilon)$ such that $|f_n(x) - f(x)| < \varepsilon$ for all $x \in [a,b]$ and all $n \geq N(\varepsilon)$.

  An infinite series $\Sigma_1^\infty f_n(x)$ converges uniformly on $[a,b]$ if its sequence of partial sums converges uniformly on $[a,b]$.

  The *Cauchy criterion* is used to establish uniform convergence. We say that an infinite sequence $\{f_n(x)\}$ of functions defined on $[a,b]$ converges uniformly there if and only if, for every $\varepsilon > 0$, there exists a fixed integer $N(\varepsilon)$ such that $|f_n(x) - f_m(x)| < \varepsilon$ for all $x \in [a,b]$ and all $n, m \geq N(\varepsilon)$.

Uniform convergence is an essential hypothesis in many theorems. For example, if $\{f_n(x)\}$ is an infinite sequence of continuous functions on $[a,b]$ and if the sequence $\{f_n(x)\}$ converges uniformly to the limit function $f(x)$ there, then $f(x)$ is continuous on $[a,b]$.

Uniform convergence is also required to justify term-by-term integration. If $\{f_n(x)\}$ is a sequence of integrable functions that converge uniformly to $f(x)$ on $[a,b]$, then $f(x)$ is integrable and

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{n\to\infty} \int_a^b f_n(x)\,\mathrm{d}x.$$

As an immediate consequence, we can say that if

$$f(x) = \sum_{n=1}^{\infty} f_n(x)$$

and the convergence is uniform on the interval $[a,b]$, then

$$\int_a^b f(x)\,\mathrm{d}x = \sum_{n=1}^{\infty} \int_a^b f_n(x)\,\mathrm{d}x.$$

- *Analytic, entire, and meromorphic functions*: A *region* is a nonempty, connected, open subset of the plane. If $f(z)$ is a complex-valued function defined in the region Ω and $f(z)$ is differentiable for every $z_0 \in \Omega$, then we say that $f(z)$ is *analytic* or *holomorphic* in Ω.

  Suppose that $f_1(z)$ is analytic on Ω_1 and that $f_2(z)$ is analytic on Ω_2. If $\Omega_1 \cap \Omega_2 \neq \emptyset$ and $f_1(z) = f_2(z)$ for all $z \in \Omega_1 \cap \Omega_2$, then $f_1(z)$ and $f_2(z)$ are said to be *direct analytic continuations* of one another.

  A function that is analytic in the whole complex plane $\mathbb{C}$ is called *entire*.

  A function $f(z)$ is *meromorphic* in a region Ω if there is a set P such that (a) P has no limit point in Ω, (b) $f(z)$ is analytic in the region $\Omega \setminus P$, and (c) $f(z)$ has a pole at each point of P. Every meromorphic function in a region Ω is a quotient of two functions which are analytic in Ω.

2.2 The Method of Successive Substitution

Consider the integral equation

$$\phi(x) = f(x) + \lambda \int_a^b K(x,t)\,\phi(t)\,\mathrm{d}t,$$

where we assume as usual that $f(x)$ is continuous on the interval $[a,b]$ and that $K(x,t)$ is complex-valued and continuous on the square $Q(a,b)$. If this equation has a solution $\phi(x)$, then the equation itself provides a representation for it. Without any

preconditions other than its integrability, $\phi(x)$ can be substituted into the integrand, thereby producing yet another, albeit more complicated, representation for $\phi(x)$ in terms of itself. More precisely, if we replace x by t and t by s in the integral equation, then a direct substitution yields

$$\begin{aligned}\phi(x) &= f(x) + \lambda \int_{t=a}^{b} K(x,t) \left[f(t) + \lambda \int_{s=a}^{b} K(t,s)\,\phi(s)\,\mathrm{d}s \right] \mathrm{d}t \\ &= f(x) + \lambda \int_{a}^{b} K(x,t)\,f(t)\,\mathrm{d}t + \lambda^2 \int_{t=a}^{b} \int_{s=a}^{b} K(x,t)\,K(t,s)\,\phi(s)\,\mathrm{d}s\,\mathrm{d}t.\end{aligned}$$

After interchanging the order of integration in the last integral and replacing the dummy variable s with t, we have

$$\phi(x) = f(x) + \lambda \int_{a}^{b} K(x,t)\,f(t)\,\mathrm{d}t + \lambda^2 \int_{a}^{b} K_2(x,t)\,\phi(t)\,\mathrm{d}t,$$

where we have set

$$K_2(x,t) = \int_{a}^{b} K(x,s)\,K(s,t)\,\mathrm{d}s.$$

It may seem pointless to repeat this process, since the solution $\phi(x)$ will always be represented in terms of itself, no matter how many times we repeat it. In actuality, however, successive substitution proves to be quite fruitful. Not only does the continuation of this iterative process eventually produce the solution to the equation, but it also produces an elegant representation for it that does not involve $\phi(x)$. Additional iterations lead to the general form

$$\phi(x) = f(x) + \sum_{m=1}^{n} \lambda^m \left(\int_{a}^{b} K_m(x,t)\,f(t)\,\mathrm{d}t \right) + \lambda^{n+1} \int_{a}^{b} K_{n+1}(x,t)\,\phi(t)\,\mathrm{d}t,$$

for any integer n, where initially we have set $K_1(x,t) = K(x,t)$, and then

$$K_m(x,t) = \int_{a}^{b} K_{m-1}(x,s)\,K(s,t)\,\mathrm{d}s, \tag{2.1}$$

for each $m = 2,\ldots,n$. This general form is valid for all λ.

The functions $K_m(x,t)$ are called *iterated kernels*. Each $K_m(x,t)$ is complex-valued and continuous on $Q(a,b)$. Also, $K_m(x,t)$ is bounded there for all $m \geq 2$, for if $|K(x,t)| \leq M$, then $|K_m(x,t)| \leq M^m\,(b-a)^{m-1}$.

Now let

$$\sigma_n(x) = \sum_{m=1}^{n} \lambda^{m-1} \left(\int_{a}^{b} K_m(x,t)\,f(t)\,\mathrm{d}t \right) \tag{2.2}$$

and

$$\rho_n(x) = \lambda^{n+1} \int_{a}^{b} K_{n+1}(x,t)\,\phi(t)\,\mathrm{d}t,$$

so that

$$\phi(x) = f(x) + \lambda\, \sigma_n(x) + \rho_n(x).$$

With a little effort, we show that the sequence $\{\sigma_n(x)\}$ of continuous functions converges uniformly to a continuous limit function $\sigma(x)$ on the interval $[a,b]$ and that $\phi(x) = f(x) + \lambda\, \sigma(x)$.

Given the bound on $K_m(x,t)$, each term of the sum $\sigma_n(x)$ satisfies the inequality

$$\left| \lambda^{m-1} \left(\int_a^b K_m(x,t)\, f(t)\, \mathrm{d}t \right) \right| \leq (\, |\lambda|\, M\, (b-a)\,)^{m-1}\, M\, \|f\|_1.$$

If $|\lambda|\, M\, (b-a) < 1$, then the sequence $\{\sigma_n(x)\}$ of partial sums is a Cauchy sequence. For if $\varepsilon > 0$ is arbitrary, then

$$\begin{aligned} |\sigma_n(x) - \sigma_p(x)| &\leq \left[\sum_{m=p+1}^{n} (\, |\lambda|\, M\, (b-a)\,)^{m-1} \right] M\, \|f\|_1 \\ &\leq (|\lambda|\, M\, (b-a))^p\, \frac{M \|f\|_1}{1 - |\lambda|\, M\, (b-a)} \\ &< \varepsilon \end{aligned}$$

if p is large enough. In addition, the remainder term $\rho_n(x) \to 0$ uniformly on $[a,b]$ as $n \to +\infty$, in view of the estimate $|\rho_n(x)| \leq |\lambda|\, M\, \|\phi\|_1\, (|\lambda|\, M\, (b-a))^n$.

It now follows that the sequence $\{\sigma_n(x)\}$ of continuous functions converges absolutely and uniformly on the interval $[a,b]$ to the continuous limit function

$$\sigma(x) = \sum_{m=1}^{\infty} \lambda^{m-1} \left(\int_a^b K_m(x,t)\, f(t)\, \mathrm{d}t \right),$$

provided that $|\lambda|\, M\, (b-a) < 1$. Furthermore, since term-by-term integration is permitted as a consequence of uniform convergence, we have

$$\sigma(x) = \int_a^b \left(\sum_{m=1}^{\infty} \lambda^{m-1}\, K_m(x,t) \right) f(t)\, \mathrm{d}t = \int_a^b R(x,t;\lambda) f(t)\, \mathrm{d}t,$$

where $R(x,t;\lambda)$ denotes the infinite series in parentheses. This series is known as the *Neumann series*, and it is the *resolvent kernel* of the integral equation. Its radius of convergence is at least $1/(M(b-a))$.

Recall that a function $\phi(x)$ is a *solution* to an integral equation if the integral equation is transformed into an identity by the substitution of $\phi(x)$ into it.

We shall show that $f(x) + \lambda\, \sigma(x)$ is the solution to the integral equation by showing that

$$f(x) + \lambda\, \sigma(x) = f(x) + \lambda \int_a^b K(x,t)\, (f(t) + \lambda\, \sigma(t))\, \mathrm{d}t.$$

By substituting the series expansion for $\sigma(x)$ into the integral on the right, we obtain

$$\begin{aligned}
&\lambda \int_a^b K(x,t)\,[f(t)+\lambda\,\sigma(t)]\,\mathrm{d}t \\
&\quad = \lambda \int_a^b K(x,t)\,f(t)\,\mathrm{d}t \\
&\qquad + \lambda^2 \int_{t=a}^b K(x,t) \left[\int_{s=a}^b \left(\sum_{m=1}^{\infty} \lambda^{m-1} K_m(t,s) \right) f(s)\,\mathrm{d}s \right] \mathrm{d}t \\
&\quad = \lambda \int_a^b K(x,t)\,f(t)\,\mathrm{d}t \\
&\qquad + \lambda^2 \int_{s=a}^b \sum_{m=1}^{\infty} \lambda^{m-1} \left(\int_{t=a}^b K_m(x,t)\,K(t,s)\,\mathrm{d}t \right) f(s)\,\mathrm{d}s \\
&\quad = \lambda \int_a^b K(x,t)\,f(t)\,\mathrm{d}t + \lambda \int_a^b \left(\sum_{m=1}^{\infty} \lambda^{m} K_{m+1}(x,s) \right) f(s)\,\mathrm{d}s \\
&\quad = \lambda\,\sigma(x).
\end{aligned}$$

We can now rightfully claim that $\phi(x) = f(x) + \lambda\,\sigma(x)$. Since $f(x)$ is assumed to be continuous on the interval $[a,b]$ and we have proven that $\sigma(x)$ is continuous there, it follows that $\phi(x)$ is continuous as well and that its norm $\|\phi\|_1$ is finite.

The discussion above serves to establish the following result:

Theorem 2.2.1 (Successive Substitution). *Let $f(x)$ be a complex-valued and continuous function defined on the interval $[a,b]$, and let $K(x,t)$ be a complex-valued continuous kernel defined on the square $Q(a,b)$ and also bounded there by M. Let λ be a complex parameter. If $|\lambda|\,M\,(b-a) < 1$, then the unique solution to the Fredholm integral equation*

$$\phi(x) = f(x) + \lambda \int_a^b K(x,t)\,\phi(t)\,\mathrm{d}t$$

is given by

$$\phi(x) = f(x) + \lambda \int_a^b R(x,t;\lambda)\,f(t)\,\mathrm{d}t,$$

where $R(x,t;\lambda)$ is the resolvent kernel

$$R(x,t;\lambda) = \sum_{m=1}^{\infty} \lambda^{m-1}\,K_m(x,t).$$

Section 2.2 Exercises

1. For each of the kernels $K(x,t)$ given below, determine all of the iterated kernels $K_m(x,t)$ and then form the resolvent kernel $R(x,t;\lambda)$. Note that the resolvent depends upon the endpoints of the interval $[a,b]$. For which values of λ does the resolvent series converge? Can you sum the resolvent series in closed form?

   (a) $A(x,t) = x+t$, defined on $Q(-1,+1)$
   (b) $B(x,t) = x-t$, defined on $Q(0,1)$
   (c) $C(x,t) = xt$, defined on $Q(a,b)$
   (d) $D(x,t) = (1+x)(1-t)$, defined on $Q(-1,0)$
   (e) $E(x,t) = x^2t^2$, defined on $Q(-1,+1)$
   (f) $F(x,t) = \cos x \cos t$, defined on $Q(0,\pi)$
   (g) $G(x,t) = \sin(x+t)$, defined on $Q(0,\pi)$
   (h) $H(x,t) = x\mathrm{e}^t$, defined on $Q(0,1)$
   (i) $I(x,t) = \mathrm{e}^{x-t}$, defined on $Q(0,1)$
   (j) $J(x,t) = \mathrm{e}^x \cos t$, defined on $Q(0,\pi)$

2. Suppose that the kernel $K(x,t)$ is continuous on the square $Q(a,b)$. Prove that if the kernel $K(x,t)$ is separable, then every iterated kernel $K_m(x,t)$, $m = 2,3,\ldots,$ is also separable.
3. Suppose that the kernel $K(x,t)$ is a function of the product xt, say $K(x,t) = G(xt)$. Is the iterated kernel $K_2(x,t)$ also a function of the product xt?
4. Suppose that the kernel $K(x,t)$ is a function of the difference $x-t$, say $K(x,t) = H(x-t)$. Is the iterated kernel $K_2(x,t)$ also a function of the difference $x-t$?
5. Is every kernel an iterated kernel? In other words, if the kernel $K(x,t)$ is continuous on the square $Q(a,b)$, does there exist a continuous kernel $L(x,t)$ such that
$$K(x,t) = \int_a^b L(x,s)\,L(s,t)\,\mathrm{d}s?$$
6. Let $m \geq 2$. Show that
$$K_m(x,t) = \int_a^b K_r(x,s)\,K_{m-r}(s,t)\,\mathrm{d}s$$
for every $r = 1,\ldots,m-1$.
Hint: If $m = 2$, then the conclusion holds by definition. Use an inductive argument for $m \geq 3$.

2.3 The Method of Successive Approximation

In this section, we introduce another recursive method for solving Fredholm integral equations of the second kind. There are some distinct advantages in considering another method: additional insight is gained into the recursive process, a different

proof of convergence is used, and a better result is obtained, in the sense that the resolvent series may have a larger radius of convergence.

Consider the Fredholm integral equation

$$\phi(x) = f(x) + \lambda \int_a^b K(x,t)\,\phi(t)\,\mathrm{d}t. \tag{2.3}$$

If $\lambda = 0$, then $\phi(x) = f(x)$ is the unique solution to the equation. In view of this observation, it seems reasonable to suggest that if $|\lambda| \approx 0$, then $\phi(x) \approx f(x)$, i.e., that the free term $f(x)$ serves as a reasonable zeroth-order approximation $\phi_0(x)$ to the solution $\phi(x)$. However, for computational reasons, $f(x)$ may not be the most practical choice for larger values of λ. Depending upon the form of the kernel, the choices $\phi_0(x) = 1$ or $\phi_0(x) = \mathrm{e}^x$ or some other choice might be more reasonable from the perspective of integrability.

After $\phi_0(x)$ is chosen, a first-order approximation $\phi_1(x)$ to $\phi(x)$ is furnished by substituting $\phi_0(t)$ for $\phi(t)$ into the integrand to obtain

$$\phi_1(x) = f(x) + \lambda \int_a^b K(x,t)\,\phi_0(t)\,\mathrm{d}t.$$

If $\phi_1(x) = \phi_0(x)$, then $\phi(x) = \phi_1(x)$, and the solution has been found. If the integral vanishes, then $\phi_1(x) = f(x)$. At this point, the iterative process continues as if we had originally chosen $\phi_0(x) = f(x)$.

If $\phi_1(x) \neq \phi_0(x)$, then the substitution of $\phi_1(t)$ into the integrand yields the second-order approximation

$$\phi_2(x) = f(x) + \lambda \int_a^b K(x,t)\,\phi_1(t)\,\mathrm{d}t.$$

If $\phi_2(x) = \phi_1(x)$, then $\phi(x) = \phi_2(x)$, and the solution has been found. If the integral vanishes, then $\phi_2(x) = f(x)$. Again, the iterative process continues as if we had originally chosen $\phi_0(x) = f(x)$. If $\phi_2(x) = \phi_0(x)$, then a continuation of this iterative process produces the two distinct constant subsequences, namely, $\phi_0(x) = \phi_2(x) = \phi_4(x) = \cdots$ and $\phi_1(x) = \phi_3(x) = \phi_5(x) = \cdots$. If this were to happen, then a unique solution to the integral equation obviously would not exist.

If $\phi_2(x) \neq \phi_1(x)$ or $\phi_0(x)$, then a substitution of $\phi_2(t)$ into the integrand yields the third-order approximation

$$\phi_3(x) = f(x) + \lambda \int_a^b K(x,t)\,\phi_2(t)\,\mathrm{d}t,$$

and comments similar to those made above regarding the two previous iterations can again be made.

Assuming that $\phi_n(x) \neq \phi_j(x)$ for $j = 0, \ldots, n-1$ and that $\phi(x) \neq f(x)$, then a substitution of $\phi_n(t)$ into the integrand yields the $(n+1)$st-order approximation

$$\phi_{n+1}(x) = f(x) + \lambda \int_a^b K(x,t)\, \phi_n(t)\, \mathrm{d}t.$$

Each approximant $\phi_n(x)$ has an alternate form. In Sect. 2.2, we substituted the integral equation into itself repeatedly; here, we substitute each approximant $\phi_j(x)$ into the expression for the next approximant $\phi_{j+1}(x)$ to obtain

$$\begin{aligned}\phi_{n+1}(x) = f(x) + \sum_{m=1}^{n} \lambda^m \left(\int_a^b K_m(x,t)\, f(t)\, \mathrm{d}t \right) \\ + \lambda^{n+1} \int_a^b K_{n+1}(x,t)\, \phi_0(t)\, \mathrm{d}t. \end{aligned} \tag{2.4}$$

Alternately, we have

$$\phi_{n+1}(x) = f(x) + \lambda\, \sigma_n(x) + \omega_{n+1}(x),$$

where $\sigma_n(x)$ was previously defined by (2.2), and we set

$$\omega_{n+1}(x) = \lambda^{n+1} \int_a^b K_{n+1}(x,t)\, \phi_0(t)\, \mathrm{d}t.$$

In Sect. 2.2, we showed that the sequence $\sigma_n(x)$ converges uniformly to the function $\sigma(x)$ on the interval $[a,b]$, provided that $|\lambda|$ was small enough relative to the size of the kernel. Here, we reprove this result under a different condition on λ.

An application of the Cauchy–Schwarz inequality to the definition of the iterated kernel yields

$$|K_m(x,t)|^2 \leq \left(\int_a^b |K_{m-1}(x,s)|^2\, \mathrm{d}s \right) \left(\int_a^b |K(s,t)|^2\, \mathrm{d}s \right).$$

Integrating this inequality with respect to t yields

$$\int_a^b |K_m(x,t)|^2\, \mathrm{d}t \leq \left(\int_a^b |K_{m-1}(x,s)|^2\, \mathrm{d}s \right) \left(\int_a^b \int_a^b |K(s,t)|^2\, \mathrm{d}s\, \mathrm{d}t \right),$$

for each fixed $x \in [a,b]$, or more simply

$$\kappa_m(x) \leq \kappa_{m-1}(x)\, \|K\|_2^2,$$

where we have set

$$\kappa_m(x) = \int_a^b |K_m(x,t)|^2\, \mathrm{d}t$$

and

$$\|K\|_2 = \left(\int_a^b \int_a^b |K(x,t)|^2 \, \mathrm{d}x \, \mathrm{d}t \right)^{1/2}.$$

Recursively, we easily obtain the estimate

$$\kappa_m(x) \le \kappa_1(x) \, \|K\|_2^{2m-2}.$$

Another application of the Cauchy–Schwarz inequality to the integrals in the sum $\sigma_n(x)$ provides the estimate

$$\begin{aligned} \left| \int_a^b K_m(x,t) \, f(t) \, \mathrm{d}t \right|^2 &\le \left(\int_a^b |K_m(x,t)|^2 \, \mathrm{d}t \right) \left(\int_a^b |f(t)|^2 \, \mathrm{d}t \right) \\ &= \kappa_m(x) \, \|f\|_2^2 \\ &\le \kappa_1(x) \, \|f\|_2^2 \, \|K\|_2^{2m-2}. \end{aligned}$$

Hence, each term in the sum $\sigma_n(x)$ can be estimated by the inequality

$$\left| \lambda^m \int_a^b K_m(x,t) \, f(t) \, \mathrm{d}t \right| \le \frac{\sqrt{\kappa_1(x)} \, \|f\|_2}{\|K\|_2} \, (|\lambda| \, \|K\|_2)^m, \tag{2.5}$$

which is valid for each fixed $x \in [a,b]$.

It follows from this estimate that the sequence $\sigma_n(x)$ converges absolutely and uniformly to a unique limit $\sigma(x)$ on the interval $[a,b]$ whenever $|\lambda| \, \|K\|_2 < 1$, since it is dominated by a convergent geometric series of positive terms.

By using a similar estimate, we also see that

$$|\omega_{n+1}(x)| \le \frac{\sqrt{\kappa_1(x)} \, \|\phi_0\|_2}{\|K\|_2} \, (|\lambda| \, \|K\|_2)^{n+1} \to 0$$

as $n \to +\infty$.

The results of the last two paragraphs imply that $\phi(x) = f(x) + \lambda \, \sigma(x)$.

The proof that $\phi(x)$ is unique proceeds by contradiction. Suppose that there were two distinct solutions to Eq. (2.3), say $\phi(x)$ and $\tilde{\phi}(x)$. If we set $\delta(x) = \phi(x) - \tilde{\phi}(x)$, then $\delta(x)$ satisfies the homogeneous integral equation

$$\delta(x) = \lambda \int_a^b K(x,t) \, \delta(t) \, \mathrm{d}t.$$

If $\delta(x) \neq 0$, then it would be an eigenfunction of the kernel corresponding to the eigenvalue λ. Showing that $\delta(x) \equiv 0$ not only proves that $\phi(x)$ is unique, but it also shows that the kernel has no eigenvalues smaller than $1/\|K\|_2$. An application of the Cauchy–Schwarz inequality gives

$$|\delta(x)|^2 \le |\lambda|^2 \left(\int_a^b |K(x,t)|^2 \, \mathrm{d}t \right) \left(\int_a^b |\delta(t)|^2 \, \mathrm{d}t \right),$$

which can be integrated with respect to x yielding

$$\left(1 - |\lambda|^2 \|K\|_2^2\right) \int_a^b |\delta(x)|^2 \, \mathrm{d}x \le 0.$$

The restriction $|\lambda| \, \|K\|_2 < 1$ implies that $\int_a^b |\delta(x)|^2 \, \mathrm{d}x = 0$.

The discussion above serves to establish the following result:

Theorem 2.3.1 (Successive Approximation). *Let λ be a complex parameter and let $f(x)$ be a complex-valued continuous function defined on the interval $[a,b]$. Let $K(x,t)$ be a complex-valued continuous kernel defined on the square $Q(a,b)$ with*

$$\|K\|_2 = \left(\int_a^b \int_a^b |K(x,t)|^2 \, \mathrm{d}x \, \mathrm{d}t \right)^{1/2}.$$

If $|\lambda| \, \|K\|_2 < 1$, then the unique solution to the Fredholm integral equation

$$\phi(x) = f(x) + \lambda \int_a^b K(x,t) \, \phi(t) \, \mathrm{d}t$$

is given by

$$\phi(x) = f(x) + \lambda \int_a^b R(x,t;\lambda) \, f(t) \, \mathrm{d}t,$$

where $R(x,t;\lambda)$ is the resolvent kernel

$$R(x,t;\lambda) = \sum_{m=1}^{\infty} \lambda^{m-1} \, K_m(x,t).$$

The estimate for the radius of convergence in the Theorem of Successive Approximation is larger than the corresponding estimate in the Theorem of Successive Substitution. Indeed, $\|K\|_2 \le M(b-a)$, with equality holding only if $K(x,t) \equiv M$. Hence,

$$\frac{1}{M(b-a)} \le \frac{1}{\|K\|_2}.$$

For example, consider the kernel $K(x,t) = x^{10} t^{10}$ on the square $Q(0,1)$. For this kernel, we have $\|K\|_2 = \frac{1}{21}$ and $M(b-a) = 1$.

The inequalities established within the proof of the theorem may be used again to estimate the magnitude of the error incurred by using the approximant $\phi_n(x)$ to estimate $\phi(x)$. For each $x \in [a,b]$, we have

$$
\begin{aligned}
|\phi(x)-\phi_n(x)| &= \left| \sum_{m=n+1}^{\infty} \lambda^m \int_a^b K_m(x,t)\, f(t)\, \mathrm{d}t \right| \\
&\le \frac{\sqrt{\kappa_1(x)}\, \|f\|_2}{\|K\|_2} \sum_{m=n+1}^{\infty} (|\lambda|\, \|K\|_2)^m \\
&\le \frac{\|\sqrt{\kappa_1}\|_\infty \|f\|_2}{\|K\|_2 (1-|\lambda|\, \|K\|_2)} (|\lambda|\, \|K\|_2)^{n+1}.
\end{aligned}
$$

This inequality shows that if n is large enough, then the difference $|\phi(x)-\phi_n(x)|$ is uniformly small on the interval $[a,b]$. However, the rate of convergence may be quite slow if $|\lambda|\, \|K\|_2$ is close to 1. The implication here is that a large number of approximants may need to be computed, thereby decreasing the efficiency and convenience of the method.

Illustrative Examples

- *Example 1*: If it should happen that an iterated kernel $K_m(x,t) \equiv 0$ on the square $Q(a,b)$, then the resolvent kernel reduces to a polynomial in λ, and the kernel $K(x,t)$ will have no eigenvalues.

  In particular, if $K_2(x,t) \equiv 0$, then $R(x,t;\lambda) = K(x,t)$, and the kernel $K(x,t)$ is said to be *orthogonal to itself*. It is easy to invent such kernels in terms of trigonometric functions. For example, if $K(x,t) = \sin(x+2t)$, then $K_2(x,t) = 0$. Hence, the solution to the integral equation

$$\phi(x) = x + \int_0^{2\pi} \sin(x+2t)\, \phi(t)\, \mathrm{d}t$$

  is

$$\phi(x) = x + \int_0^{2\pi} \sin(x+2t)\, t\, \mathrm{d}t = x - \pi \cos x.$$

- *Example 2*: Consider the separable kernel

$$K(x,t) = \cos(x+t) = \cos x \cos t - \sin x \sin t$$

  defined on the square $Q(0,\pi)$. By employing the methods in Chap. 1, we can obtain the resolvent kernel in the form

$$R(x,t;\lambda) = -\frac{D(x,t;\lambda)}{D(\lambda)} = \frac{\cos x \cos t}{1-\frac{\lambda\pi}{2}} - \frac{\sin x \sin t}{1+\frac{\lambda\pi}{2}},$$

  showing that it is a meromorphic function of λ with simple poles at $\pm 2/\pi$.

Alternatively, the resolvent kernel can also be easily computed in terms of the iterated kernels of $K(x,t)$. Simple integrations yield

$$K_2(x,t) = \int_0^\pi \cos(x+s)\cos(s+t)\,\mathrm{d}s = \tfrac{\pi}{2}\cos(x-t)$$

and

$$K_3(x,t) = \int_0^\pi \cos(x-s)\cos(s+t)\,\mathrm{d}s = \left(\tfrac{\pi}{2}\right)^2 \cos(x+t) = \left(\tfrac{\pi}{2}\right)^2 K(x,t).$$

A short inductive argument shows that

$$K_m(x,t) = \begin{cases} \left(\frac{\pi}{2}\right)^{m-1}\cos(x+t), & \text{if } m \text{ is odd,} \\ \left(\frac{\pi}{2}\right)^{m-1}\cos(x-t), & \text{if } m \text{ is even.} \end{cases}$$

After inserting the iterated kernels into the resolvent series and simplifying the result, we obtain the same expression for $R(x,t;\lambda)$ as above.

The Theorem of Successive Substitution guarantees that the resolvent series converges if $|\lambda| \le 1/M(b-1) < 1/\pi$; the Theorem of Successive Approximation guarantees convergence if $|\lambda| \le 1/\|K\|_2 < \sqrt{2}/\pi$. If we were to expand $R(x,t;\lambda)$ in a Taylor series in λ about the origin, then its radius of convergence would be $2/\pi$.

- *Example 3*: Suppose that we need to solve the Fredholm integral equation

$$\psi(x) = x^4 + \tfrac{1}{10}\int_0^1 \mathrm{e}^{xt}\,\psi(t)\,\mathrm{d}t. \tag{2.6}$$

Since the kernel $K(x,t) = \mathrm{e}^{xt}$ is not separable, the methods of Chap. 1 do not apply.

In an attempt to use one of the successive methods, we easily compute

$$K_2(x,t) = \frac{\mathrm{e}^{x+t}-1}{x+t}$$

but then quickly realize that the integration required to produce the next iterate $K_3(x,t)$ is unmanageable.

Next, in an attempt to use the Method of Successive Approximation, we choose $\psi_0(x) = x^4$ and then use integration by parts several times to compute $\psi_1(x)$ as prescribed, but then we quickly find that the integration required to produce $\psi_2(x)$ is unmanageable.

In situations like this, *it may be advantageous to approximate the kernel* in our attempt to approximate the solution. Recalling that

$$\mathrm{e}^{xt} = \sum_{m=0}^{\infty} \frac{1}{m!}\,x^m t^m,$$

we truncate this series and consider the Fredholm integral equation

$$\phi(x) = x^4 + \tfrac{1}{10} \int_0^1 K(x,t)\,\phi(t)\,\mathrm{d}t \tag{2.7}$$

with the separable kernel

$$K(x,t) = 1 + xt + \frac{x^2t^2}{2!} + \frac{x^3t^3}{3!} + \frac{x^4t^4}{4!}$$

defined on the square $Q(0,1)$. All five eigenvalues of the kernel are real and positive, the smallest of which is $\lambda_1 \approx 0.739241$. Since $\frac{1}{10}$ is therefore a regular value, the integral equation has a unique solution by the First Fredholm Theorem.

A direct manual computation of the resolvent kernel is impractical since each of the iterated kernels $K_m(x,t)$ consists of 25 terms of the form $x^i y^j$. Instead, we prefer to determine a suitable approximation to the solution by using the Method of Successive Approximation. Since

$$\|K\|_2 = \frac{\sqrt{841,676,993}}{15120\sqrt{2}},$$

the Theorem of Successive Approximation guarantees that the resolvent series converges if $|\lambda| \le 1/\|K\|_2 \approx 0.737045$, a value that is remarkably close to λ_1. (The Theorem of Successive Substitution guarantees that the resolvent series converges if $|\lambda| \le 1/M(b-a) = \frac{24}{65} \approx 0.369231$.)

With the uniform convergence of the sequence $\{\phi_n(x)\}$ to the solution $\phi(x)$ now guaranteed, we can begin to compute it. Since $\lambda = \frac{1}{10}$ is small, we choose $\phi_0(x) = x^4$. In order to demonstrate the efficacy of the method to produce a convergent sequence, we list the first six approximants in the form

$$\phi_n(x) = x^4 + \tfrac{1}{10}\,\tau_n(x).$$

The coefficients of the polynomials $\tau_n(x)$ have been computed accurately here to six decimal places but not rounded up or down:

$$\begin{aligned}
\tau_1(x) &= 0.200000 + 0.166666x + 0.071428x^2 + 0.020833x^3 + 0.004629x^4\\
\tau_2(x) &= 0.231327 + 0.184501x + 0.077766x^2 + 0.022479x^3 + 0.004967x^4\\
\tau_3(x) &= 0.235611 + 0.186859x + 0.078590x^2 + 0.022692x^3 + 0.005010x^4\\
\tau_4(x) &= 0.236191 + 0.187178x + 0.078701x^2 + 0.022720x^3 + 0.005016x^4\\
\tau_5(x) &= 0.236269 + 0.187221x + 0.078717x^2 + 0.022724x^3 + 0.005017x^4\\
\tau_6(x) &= 0.236280 + 0.187226x + 0.078719x^2 + 0.022725x^3 + 0.005017x^4
\end{aligned}$$

The exact solution is given by $\phi(x) = x^4 + \frac{1}{10}\,\tau(x)$, where

$$\tau(x) = 0.236282 + 0.187227\,x + 0.078719\,x^2 + 0.022725\,x^3 + 0.005017\,x^4.$$

The rate of convergence is quite satisfactory. Indeed, the coefficients of $\tau(x)$ differ from those of $\tau_6(x)$ by at most 0.000002!

The degree of accuracy on display here could have been predicted by the error estimate given above. Since $\lambda = \frac{1}{10}$, $\|f\|_2 = \frac{1}{3}$, $\|K\|_2 = 1.3567691$, and $\|\sqrt{\kappa_1}\|_\infty = 1.7851749$, the error in using $\phi_6(x)$ to approximate $\phi(x)$ is no more than $|\phi(x) - \phi_6(x)| \le \frac{1}{10}(0.0000043)$.

To grade the success in approximating the solution $\psi(x)$ of Eq. (2.6) with the solution $\phi(x)$ of Eq. (2.7), it is reasonable to estimate the norm $\|\psi - \phi\|_\infty$.

An application of the triangle inequality yields

$$|\psi(x) - \phi(x)| \le \tfrac{1}{10} \left| \int_0^1 \mathrm{e}^{xt} \left(\psi(t) - \phi(t)\right) \mathrm{d}t \right| + \tfrac{1}{10} \left| \int_0^1 \left(\mathrm{e}^{xt} - K(x,t)\right) \phi(t)\, \mathrm{d}t \right|.$$

For each fixed $x \in [0,1]$, we apply the Cauchy–Schwarz inequality to each of these integrals to obtain

$$|\psi(x) - \phi(x)| \le \tfrac{1}{10} \|\mathrm{e}^{xt}\|_2 \cdot \|\psi - \phi\|_2 + \tfrac{1}{10} \|\mathrm{e}^{xt} - K(x,t)\|_2 \cdot \|\phi\|_2.$$

On the square $Q(0,1)$, we have

$$\|\mathrm{e}^{xt}\|_2 < \tfrac{9}{5}, \quad \text{and} \quad \|\mathrm{e}^{xt} - K(x,t)\|_2 < \tfrac{1}{338}$$

independently of x, so that the last inequality becomes

$$|\psi(x) - \phi(x)| \le \tfrac{9}{50} \|\psi - \phi\|_2 + \tfrac{1}{3380} \|\phi\|_2.$$

Taking the supremum over the left side of this inequality and replacing $\|\psi - \phi\|_2$ with the greater norm $\|\psi - \phi\|_\infty$ gives

$$\|\psi - \phi\|_\infty \le \tfrac{9}{50} \|\psi - \phi\|_\infty + \tfrac{1}{3380} \|\phi\|_2.$$

Finally, since

$$\|\phi\|_2 \le \|\phi_6\|_2 + \|\phi - \phi_6\|_2 \le \|\phi_6\|_2 + \|\phi - \phi_6\|_\infty < \tfrac{3}{8},$$

we obtain the uniform estimate

$$\|\psi - \phi\|_\infty \le \tfrac{50}{41} \cdot \tfrac{1}{3380} \cdot \tfrac{3}{8} = \tfrac{15}{110864} < \tfrac{1}{7390}.$$

Also, given the error estimate for $\|\phi - \phi_6\|_\infty$ that was established above, we have

$$\|\psi - \phi_6\|_\infty \le \|\psi - \phi\|_\infty + \|\phi - \phi_6\|_\infty < \tfrac{1}{7366}.$$

Practically speaking, the graphs of $\psi(x)$ and $\phi_6(x)$ are "within the ink" of each other, since their values agree so closely everywhere within the interval $[0,1]$.

Section 2.3 Exercises

1. Solve the nonhomogeneous Fredholm integral equation

$$\phi(x) = ax + \lambda \int_0^1 xt\,\phi(t)\,\mathrm{d}t$$

in two ways.

(a) Use the techniques in Chap. 1 to show that the unique solution to this equation is given by

$$\phi(x) = \frac{ax}{1 - \frac{\lambda}{3}}$$

if $\lambda \neq 3$.

(b) Use the Method of Successive Approximation, with the initial choice $\phi_0(x) = ax$, to compute $\phi_n(x)$, for every $n \geq 1$. Show that the limit $\phi(t) = \lim_{n\to\infty} \phi_n(x)$ exists. The solution obtained here should agree with the solution obtained above within the circle of convergence.

2. Suppose that the kernel $K(x,t)$ is continuous on the square $Q(a,b)$.

(a) If $K(x,x) \equiv 0$ on $[a,b]$, is it true that $K_2(x,x) \equiv 0$ on $[a,b]$?
(b) If $K(x,t) = K(t,x)$, is it true that $K_m(x,t) = K_m(t,x)$?
(c) If $K(x,t) = -K(t,x)$, is it true that $K_m(x,t) = (-1)^m K_m(t,x)$?

3. Suppose that the kernels $K(x,t)$ and $L(x,t)$ are continuous on the square $Q(a,b)$. If $K_2(x,t) = L_2(x,t)$, is it true that $K(x,t) = L(x,t)$?
4. Suppose that the kernel $K(x,t)$ is continuous on the symmetric square $Q(-a,+a)$ and that

$$K(x,y) = K(-x,y) = K(x,-y) = K(-x,-y).$$

Does $K_m(x,t)$ have the same property for all $m \geq 2$?
5. Suppose that the kernel $K(x,t)$ is continuous on the square $Q(a,b)$ and that $K(x,t) \equiv 0$ for all $a \leq x \leq t \leq b$, i.e., the kernel vanishes on and above the diagonal of the square. Does $K_2(x,t)$ have the same property?
6. Define

$$K(x,t) = \begin{cases} t(1-x) & \text{if } t \leq x \\ x(1-t) & \text{if } x \leq t \end{cases}$$

on the square $Q(0,1)$. This kernel vanishes on the boundary of the square, and $K(x,x) = x(1-x)$ on its diagonal. Compute $K_2(x,t)$.
7. Define $K(x,t) = |x-t|$ on the square $Q(0,a)$. Compute $K_2(x,t)$.

8. By using its series representation, show that the resolvent satisfies the equations

$$R(x,t;\lambda) = K(x,t) + \lambda \int_a^b K(x,s)\, R(s,t;\lambda)\, \mathrm{d}s$$

and

$$R(x,t;\lambda) = K(x,t) + \lambda \int_a^b R(x,s;\lambda)\, K(s,t)\, \mathrm{d}s.$$

9. By using its series representation, show that the resolvent satisfies the integro-differential equation

$$\frac{\partial}{\partial \lambda} R(x,t;\lambda) = \int_a^b R(x,s;\lambda)\, R(s,t;\lambda)\, \mathrm{d}s.$$

10. Let $K(x,t)$ be a continuous kernel defined on $Q(a,b)$, and let

$$\|K\|_2 = \left(\int_a^b \int_a^b |K(x,t)|^2 \, \mathrm{d}x \, \mathrm{d}t \right)^{1/2}.$$

Assume that λ is small enough so that

$$\|R_\lambda\|_2 = \left(\int_a^b \int_a^b |R(x,t;\lambda)|^2 \, \mathrm{d}x \, \mathrm{d}t \right)^{1/2} < +\infty.$$

Estimate $\|R_\lambda\|_2$ in terms of $|\lambda|$ and $\|K\|_2$. For which values of λ is your estimate valid?

11. Consider the integral equation

$$\psi(x) = x + \tfrac{1}{2} \int_0^1 \cosh(xt)\, \psi(t)\, \mathrm{d}t$$

and assume that $\frac{1}{2}$ is not an eigenvalue of the kernel. Choose $\psi_0(x) = x$ and compute $\psi_1(x)$, $\psi_2(x)$, and $\psi_3(x)$. (Recall that $\cosh x = (\mathrm{e}^x + \mathrm{e}^{-x})/2$.) Next, approximate the kernel by

$$\cosh(xt) \approx 1 + \tfrac{1}{2} x^2 t^2 + \tfrac{1}{24} x^4 t^4$$

and consider the integral equation

$$\phi(x) = x + \tfrac{1}{2} \int_0^1 \left(1 + \tfrac{1}{2} x^2 t^2 + \tfrac{1}{24} x^4 t^4\right) \phi(t)\, \mathrm{d}t.$$

Choose $\phi_0(x) = x$, and compute $\phi_1(x)$, $\phi_2(x)$, and $\phi_3(x)$. Compare your results, and estimate the accuracy of your approximations by following the techniques in Example 3.

2.4 The Fredholm Theorems

In Sect. 1.3, we proved the Fredholm theorems for the integral equation

$$\phi(x) = f(x) + \lambda \int_a^b K(x,t)\,\phi(t)\,\mathrm{d}t, \tag{2.8}$$

where λ is an arbitrary complex parameter and $K(x,t)$ is a separable kernel. Then, in Sect. 2.3, we showed that the Method of Successive Approximations can be successfully applied to solve this equation if $K(x,t)$ is a general kernel, provided that $|\lambda| < 1/\|K\|_2$. In this section, we combine the techniques of those two sections in order to establish the Fredholm theorems if $K(x,t)$ is a general kernel and λ is arbitrary.

An ingenious idea due to Schmidt enables the discussion. His idea is to show that the solution $\phi(x)$ to Eq. (2.8) satisfies *two* integral equations, one whose kernel is small and another one whose kernel is separable. The techniques delineated in Sects. 1.3 and 2.3 can then be applied to these equations.

Some assumptions are in order. The complex parameter λ is assumed at first to belong to the closed disk $\Delta_\rho = \{\lambda : |\lambda| \le \rho\}$, where ρ is arbitrarily large but fixed. As usual, the free term $f(x)$ is assumed to be complex-valued and continuous on the interval $[a,b]$, and the kernel $K(x,t)$ is assumed to be complex-valued and continuous on the square $Q(a,b)$.

The discussion begins with a decomposition of the kernel $K(x,t)$.

By virtue of the well-known Weierstrass approximation theorem, the kernel $K(x,t)$ can be decomposed into a sum of two complex-valued and continuous kernels as

$$K(x,t) = K_{\mathrm{sep}}(x,t) + K_\varepsilon(x,t).$$

The kernel $K_{\mathrm{sep}}(x,t)$ can be chosen to be a separable polynomial in the variables x and t in the form

$$K_{\mathrm{sep}}(x,t) = \sum_{i=1}^{n} a_i(x)\,\overline{b_i(t)},$$

where each $a_i(x)$ and $b_i(t)$ is complex-valued and continuous on $[a,b]$, and each of the sets $\{a_i(x)\}$ and $\{b_i(t)\}$ is linearly independent on the interval $[a,b]$. The kernel $K_\varepsilon(x,t)$ can be chosen so that its norm

$$\|K_\varepsilon\|_2 = \left(\int_a^b \int_a^b |K_\varepsilon(x,t)|^2\,\mathrm{d}x\,\mathrm{d}t \right)^{1/2}$$

is arbitrarily small. In particular, if we require that $\|K_\varepsilon\|_2 < \varepsilon$, we choose $\rho = 1/\varepsilon$, so that if $|\lambda| \le \rho$, then $|\lambda| < 1/\|K_\varepsilon\|_2$.

Given this decomposition of the kernel, we may write Eq. (2.8) in the form

$$\phi(x) = f(x) + \lambda \int_a^b K_{\mathrm{sep}}(x,t)\,\phi(t)\,\mathrm{d}t + \lambda \int_a^b K_\varepsilon(x,t)\,\phi(t)\,\mathrm{d}t.$$

Equivalently, we have

$$\phi(x) = F(x;\lambda) + \lambda \int_a^b K_\varepsilon(x,t)\,\phi(t)\,\mathrm{d}t, \tag{2.9}$$

which is the first of the two integral equations satisfied by $\phi(x)$, where

$$F(x;\lambda) = f(x) + \lambda \int_a^b K_{\mathrm{sep}}(x,t)\,\phi(t)\,\mathrm{d}t.$$

The continuity of $F(x;\lambda)$ on the interval $[a,b]$ is due to the integrability of $\phi(t)$. At least for $|\lambda| < 1/\|K\|_2$, the solution $\phi(x)$ that exists by virtue of the Theorem of Successive Approximation is continuous on the interval $[a,b]$.

Now consider the equation

$$\psi(x) = g(x) + \lambda \int_a^b K_\varepsilon(x,t)\,\psi(t)\,\mathrm{d}t,$$

where $g(x)$ is assumed to be complex-valued and continuous on the interval $[a,b]$. As a consequence of the Theorem of Successive Approximation, the solution $\psi(x)$ to this equation may be expressed in the form

$$\psi(x) = g(x) + \lambda \int_a^b R_\varepsilon(x,t;\lambda)\,g(t)\,\mathrm{d}t,$$

if $|\lambda| < 1/\|K_\varepsilon\|_2$. In particular, this conclusion holds for all values of $\lambda \in \Delta_\rho$, since we chose $K_\varepsilon(x,t)$ so that $\rho < 1/\|K_\varepsilon\|_2$. As a consequence of this choice, the resolvent series

$$R_\varepsilon(x,t;\lambda) = \sum_{m=1}^{\infty} \lambda^{m-1} K_{\varepsilon m}(x,t), \tag{2.10}$$

generated from the iterates of $K_\varepsilon(x,t)$, is an analytic function of λ in Δ_ρ and is a continuous function of x and t on the square $Q(a,b)$ for each $\lambda \in \Delta_\rho$.

Whenever $F(x;\lambda)$ is continuous, the solution to Eq. (2.9) can be displayed as

$$\phi(x) = F(x;\lambda) + \lambda \int_a^b R_\varepsilon(x,t;\lambda)\,F(t;\lambda)\,\mathrm{d}t.$$

After replacing $F(x;\lambda)$ by its definition in this equation, we obtain

$$\phi(x) = f_\varepsilon(x;\lambda) + \lambda \int_a^b G_\varepsilon(x,t;\lambda)\,\phi(t)\,\mathrm{d}t, \tag{2.11}$$

where

$$f_\varepsilon(x;\lambda) = f(x) + \lambda \int_a^b R_\varepsilon(x,t;\lambda)\,f(t)\,\mathrm{d}t \tag{2.12}$$

and

$$G_\varepsilon(x,t;\lambda) = K_{\text{sep}}(x,t) + \lambda \int_a^b R_\varepsilon(x,u;\lambda)\, K_{\text{sep}}(u,t)\, \mathrm{d}u. \tag{2.13}$$

Equation (2.11) is the second integral equation that is satisfied by the solution $\phi(x)$ to Eq. (2.8). The free term $f_\varepsilon(x;\lambda)$ is an analytic function of λ in Δ_ρ, is complex-valued and continuous on the interval $[a,b]$, and exhibits an adjustment to the original free term $f(x)$ that is based on the choice of $K_\varepsilon(x,t)$. The kernel $G_\varepsilon(x,t;\lambda)$ is an analytic function of λ on Δ_ρ and is complex-valued and continuous on the square $Q(a,b)$. It reflects an adjustment to the original kernel $K(x,t)$ that is based upon its decomposition.

It is important to note that $G_\varepsilon(x,t;\lambda)$ is separable. To see this, observe that

$$\begin{aligned}
\int_a^b R_\varepsilon(x,u;\lambda)\, K_{\text{sep}}(u,t)\, \mathrm{d}u &= \int_a^b R_\varepsilon(x,u;\lambda) \left(\sum_{i=1}^n a_i(u)\, \overline{b_i(t)} \right) \mathrm{d}u \\
&= \sum_{i=1}^n \left(\int_a^b R_\varepsilon(x,u;\lambda)\, a_i(u)\, \mathrm{d}u \right) \overline{b_i(t)} \\
&= \sum_{i=1}^n A_{\varepsilon i}(x;\lambda)\, \overline{b_i(t)},
\end{aligned}$$

where

$$\begin{aligned}
A_{\varepsilon i}(x;\lambda) &= \int_a^b R_\varepsilon(x,u;\lambda)\, a_i(u)\, \mathrm{d}u \\
&= \int_a^b \left(\sum_{m=1}^\infty \lambda^{m-1} K_{\varepsilon m}(x,u) \right) a_i(u)\, \mathrm{d}u \\
&= \sum_{m=1}^\infty \lambda^{m-1} \left(\int_a^b K_{\varepsilon m}(x,u)\, a_i(u)\, \mathrm{d}u \right).
\end{aligned} \tag{2.14}$$

Term-by-term integration is permitted here, since the resolvent series converges absolutely and uniformly on the interval $[a,b]$ for every $\lambda \in \Delta_\rho$. Each $A_{\varepsilon i}(x;\lambda)$ is an analytic function of λ on Δ_ρ and is continuous on the interval $[a,b]$. With these notational conventions, the kernel $G_\varepsilon(x,t;\lambda)$ can be explicitly displayed in the separable form

$$G_\varepsilon(x,t;\lambda) = \sum_{i=1}^n \left[a_i(x) + \lambda\, A_{\varepsilon i}(x;\lambda) \right] \overline{b_i(t)}. \tag{2.15}$$

Equation (2.11) can now be treated with the techniques that were delineated in Sect. 1.3 for integral equations with separable kernels, although some differences will arise due to the fact that its kernel depends upon λ. The analysis of Eq. (2.11) begins by replacing $G_\varepsilon(x,t;\lambda)$ in with its representation (2.15) to obtain

$$\phi(x) = f_\varepsilon(x;\lambda) + \lambda \sum_{i=1}^{n} c_i(\lambda)\,[a_i(x) + \lambda A_{\varepsilon i}(x;\lambda)]\,, \tag{2.16}$$

where we have set

$$c_i(\lambda) = \int_a^b \phi(t)\,\overline{b_i(t)}\,\mathrm{d}t.$$

Every solution to Eq. (2.11) assumes this form. It remains to determine the coefficients $c_i(\lambda)$. If we replace x by t, change the index of summation from i to j, multiply both sides of Eq. (2.16) by $\overline{b_i(t)}$, and then integrate the resulting expression from a to b, then we obtain the system of linear equations

$$c_i(\lambda) = f_i(\lambda) + \lambda \sum_{j=1}^{n} c_j(\lambda)\,a_{ij}(\lambda) \tag{2.17}$$

for $i = 1,\ldots,n$, where we have set

$$f_i(\lambda) = \int_a^b f_\varepsilon(t;\lambda)\,\overline{b_i(t)}\,\mathrm{d}t$$

and

$$a_{ij}(\lambda) = \int_a^b [a_j(t) + \lambda A_{\varepsilon j}(t;\lambda)]\,\overline{b_i(t)}\,\mathrm{d}t. \tag{2.18}$$

Since the definitions of $f_i(\lambda)$ and $a_{ij}(\lambda)$ depend upon the resolvent kernel $R_\varepsilon(x,t;\lambda)$, they both are analytic functions of λ in Δ_ρ.

The linear system (2.17) can also be written in matrix form as

$$(\mathbf{I} - \lambda\,\mathbf{A}(\lambda))\,\mathbf{c}(\lambda) = \mathbf{f}(\lambda). \tag{2.19}$$

The Fredholm determinant $D_\rho(\lambda) = \det(\mathbf{I} - \lambda\,\mathbf{A}(\lambda))$ is an analytic function of λ in the closed disk Δ_ρ. An analytic function defined on an open neighborhood of a closed disk D in the complex plane that vanishes on an infinite set with a limit point in D vanishes identically on the neighborhood of D. Since $D_\rho(0) = 1$, $D_\rho(\lambda)$ can have only a finite number of zeroes in Δ_ρ. In Sect. 1.3, the matrix $\mathbf{A}$, corresponding to the matrix $\mathbf{A}(\lambda)$ here, was a matrix of constants, and the Fredholm determinant $D(\lambda) = \det(\mathbf{I} - \lambda\,\mathbf{A})$ was a polynomial of degree n with exactly n zeroes. However, since $D_\rho(\lambda)$ is not necessarily a polynomial, it may have more or less than n zeroes, but in any case, the number of zeroes is finite.

The values of $\lambda \in \Delta_\rho$ for which $D_\rho(\lambda) \neq 0$ are called *regular* values of the kernel $K(x,t)$. If λ is a regular value, then the unique solution $\phi(x)$ to the inhomogeneous integral equation (2.8) can ultimately be expressed as an integral with a resolvent kernel as stated below in the First Fredholm Theorem.

On the other hand, the values of λ for which $D_\rho(\lambda) = 0$ are called *eigenvalues of the kernel* $K(x,t)$. If λ is an eigenvalue, then the homogeneous case of the integral

equation (2.8) will have nontrivial solutions called *eigenfunctions of the kernel*; the inhomogeneous case may or may not have solutions, depending upon additional considerations.

The analysis of the linear system (2.19) proceeds in remarkedly different ways, depending upon the value of the Fredholm determinant $D_\rho(\lambda)$. The linear system, and hence the integral equation (2.8), will have exactly one solution if the determinant does not vanish. Otherwise, the linear system will have either an infinite number of solutions or no solutions at all. Both possibilities may occur, and each of them has implications for our investigation.

We consider two cases:

Case I: $D_\rho(\lambda) = \det(\mathbf{I} - \lambda\,\mathbf{A}(\lambda)) \neq 0$.

In this case, the linear system has the unique solution

$$\begin{aligned}\mathbf{c}(\lambda) &= (\mathbf{I} - \lambda\,\mathbf{A}(\lambda))^{-1}\mathbf{f}(\lambda)\\ &= \frac{1}{D_\rho(\lambda)}\,\mathrm{adj}(\mathbf{I} - \lambda\,\mathbf{A}(\lambda))\,\mathbf{f}(\lambda),\end{aligned} \tag{2.20}$$

where $\mathrm{adj}(\mathbf{I} - \lambda\,\mathbf{A}(\lambda)) = (D_{ji}(\lambda))$ is the transpose of the matrix of cofactors from $\mathbf{I} - \lambda\,\mathbf{A}(\lambda)$. Each coefficient $c_i(\lambda)$ can be represented in the form

$$c_i(\lambda) = \frac{1}{D_\rho(\lambda)} \sum_{j=1}^{n} D_{ji}(\lambda)\, f_j(\lambda).$$

Note that each $D_{ji}(\lambda)$ is an analytic function of λ in the closed disk Δ_ρ.

We consider two possibilities:

1. If $\mathbf{f}(\lambda) = \mathbf{0}$, then $\mathbf{c}(\lambda) = \mathbf{0}$, i.e., $c_i(\lambda) = 0$ for all $i = 1, \ldots, n$. Since every solution to Eq. (2.11) has the general form (2.16), it follows that $\phi(x) = f_\varepsilon(x;\lambda)$. In particular, if $f(x) \equiv 0$, then $f_\varepsilon(x;\lambda) \equiv 0$ on $[a,b]$, so that the unique solution to Eq. (2.11) is $\phi(x) \equiv 0$. It is possible for $\mathbf{f}(\lambda) = \mathbf{0}$ with $f_\varepsilon(x;\lambda) \neq 0$ on $[a,b]$. For instance, if each $b_i(t)$ happens to be orthogonal to $f(t)$ and the iterates of $K_\varepsilon(t,u)$, then $\mathbf{f}(\lambda) = \mathbf{0}$.
2. If $\mathbf{f}(\lambda) \neq \mathbf{0}$, then $\mathbf{c}(\lambda) \neq \mathbf{0}$, i.e., $c_i(\lambda) \neq 0$ for at least one subscript i. After substituting these values for $c_i(\lambda)$ into Eq. (2.16), we obtain

$$\begin{aligned}\phi(x) &= f_\varepsilon(x;\lambda) + \lambda \sum_{i=1}^{n} \left(\frac{1}{D_\rho(\lambda)} \sum_{j=1}^{n} D_{ji}(\lambda)\, f_j(\lambda) \right) [a_i(x) + \lambda\, A_{\varepsilon i}(x;\lambda)]\\ &= f_\varepsilon(x;\lambda) + \lambda \int_a^b S_\varepsilon(x,t;\lambda)\, f_\varepsilon(t;\lambda)\,\mathrm{d}t,\end{aligned}$$

where we have set

$$S_\varepsilon(x,t;\lambda) = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} D_{ji}(\lambda)\,[a_i(x) + \lambda\, A_{\varepsilon i}(x;\lambda)]\,\overline{b_j(t)}}{D_\rho(\lambda)}. \tag{2.21}$$

The kernel $S_\varepsilon(x,t;\lambda)$ is a meromorphic function of λ in the closed disk Δ_ρ, in addition to being a separable function of x and t. Since $S_\varepsilon(x,t;0) = K_{\text{sep}}(x,t)$, $S_\varepsilon(x,t;\lambda)$ represents an adjustment to $K_{\text{sep}}(x,t)$ based upon the decomposition of $K(x,t)$ and can also be written as a negative quotient of determinants.

After substituting $f_\varepsilon(x;\lambda)$ into this last representation for $\phi(x)$, we finally obtain a representation for the solution to Eq. (2.11), and therefore also Eq. (2.8), in the compact form

$$\phi(x) = f(x) + \lambda \int_a^b U_\varepsilon(x,t;\lambda)\, f(t)\, \mathrm{d}t,$$

where

$$U_\varepsilon(x,t;\lambda) = R_\varepsilon(x,t;\lambda) + S_\varepsilon(x,t;\lambda) + \lambda \int_a^b S_\varepsilon(x,s;\lambda)\, R_\varepsilon(s,t;\lambda)\, \mathrm{d}s. \tag{2.22}$$

The resolvent $R_\varepsilon(x,t;\lambda)$ is relatively small, since it was obtained by iterating $K_\varepsilon(x,t)$, and the integral term is also small, since $R_\varepsilon(s,t;\lambda)$ appears in the integrand. Hence, $S_\varepsilon(x,t;\lambda)$ is the main component of $U_\varepsilon(x,t;\lambda)$.

The kernel $U_\varepsilon(x,t;\lambda)$ is a meromorphic function of λ in the closed disk Δ_ρ. Actually, $U_\varepsilon(x,t;\lambda)$ is the restriction of a function that is meromorphic in the full λ-plane. For if $\tilde{\rho}$ is arbitrary with $0 < \rho < \tilde{\rho}$, then we can construct another resolvent kernel $\tilde{U}_\varepsilon(x,t;\lambda)$ that is meromorphic in $\Delta_{\tilde{\rho}}$, so that the solution to Eq. (2.11) can also be represented in the form

$$\phi(x) = f(x) + \lambda \int_a^b \tilde{U}_\varepsilon(x,t;\lambda)\, f(t)\, \mathrm{d}t.$$

However, since $\phi(x)$ is unique, $U_\varepsilon(x,t;\lambda) = \tilde{U}_\varepsilon(x,t;\lambda)$ on a neighborhood of the origin. Thus, $\tilde{U}_\varepsilon(x,t;\lambda)$ is a meromorphic extension of $U_\varepsilon(x,t;\lambda)$. Since $\tilde{\rho}$ was assumed to be arbitrary, it follows $U_\varepsilon(x,t;\lambda)$ may be extended to the entire λ-plane, and we let $R(x,t;\lambda)$ denote this unique extension.

The discussion above serves to establish the following result:

Theorem 2.4.1 (The First Fredholm Theorem). *Let λ be a complex parameter, $f(x)$ be a complex-valued continuous function defined on the interval $[a,b]$, and $K(x,t)$ be a complex-valued continuous kernel defined on the region $Q(a,b)$.*

Then the unique solution $\phi(x)$ to the Fredholm integral equation of the second kind

$$\phi(x) = f(x) + \lambda \int_a^b K(x,t)\, \phi(t)\, \mathrm{d}t$$

has the representation

$$\phi(x) = f(x) + \lambda \int_a^b R(x,t;\lambda)\, f(t)\, \mathrm{d}t,$$

for every regular value of λ, where the unique resolvent kernel $R(x,t;\lambda)$ is a meromorphic function of λ in the complex plane $\mathbb{C}$.

Note that if λ is a regular value of the kernel, then the homogeneous integral equation has only the trivial solution $\phi(x) \equiv 0$ on the interval $[a,b]$. For if $f(x) \equiv 0$ on $[a,b]$, then $f_\varepsilon(x;\lambda) \equiv 0$, so that $\mathbf{f}(\lambda) = \mathbf{0}$, implying in turn that $\mathbf{c}(\lambda) = \mathbf{0}$. The general form (2.16) of the solution reduces to $\phi(x) \equiv 0$.

In the discussion above, it was also established that the disk Δ_ρ, ρ arbitrary, could contain only a finite number of eigenvalues of the kernel $K(x,t)$. If the set of eigenvalues of $K(x,t)$ had a finite limit point, then there would exist some value of ρ for which Δ_ρ contained an infinite number of eigenvalues. The next result follows from this observation.

Theorem 2.4.2 (The Fourth Fredholm Theorem). *Let $K(x,t)$ be a complex-valued continuous kernel defined on the square $Q(a,b)$. Let Λ_K denote the set of eigenvalues of the kernel $K(x,t)$, that is, the set of values λ for which the homogeneous Fredholm integral equation*

$$\phi(x) = \lambda \int_a^b K(x,t)\,\phi(t)\,\mathrm{d}t$$

has a nontrivial solution. Then Λ_K is at most countable, and it cannot have a finite limit point.

Case II: $D_\rho(\lambda) = \det(\mathbf{I} - \lambda\,\mathbf{A}(\lambda)) = 0$.

In this case, there are a finite number of values $\lambda \in \Delta_\rho$ for which $D_\rho(\lambda) = 0$. We consider two possibilities:

1. If $\mathbf{f}(\lambda) = \mathbf{0}$, then the linear system

$$(\mathbf{I} - \lambda\,\mathbf{A}(\lambda))\,\mathbf{c}(\lambda) = \mathbf{0} \tag{2.23}$$

has a certain number $p(\lambda)$ of nonzero linearly independent vector solutions $\mathbf{c}^{(j)}(\lambda)$, $j = 1,\ldots,p(\lambda)$, that can be written in the form

$$\mathbf{c}^{(j)}(\lambda) = \begin{pmatrix} c_1^{(j)}(\lambda) \\ \vdots \\ c_n^{(j)}(\lambda) \end{pmatrix}.$$

Insert these values of $c_i^{(j)}(\lambda)$ into Eq. (2.16) to obtain the solution to Eq. (2.8). If $f(x) \equiv 0$ on $[a,b]$, then $f_\varepsilon(x;\lambda) \equiv 0$. For $j = 1,\ldots,p(\lambda)$, the general form of the solution to the homogeneous counterpart to the integral equation (2.8) reduces to

$$\begin{aligned} \phi_j^{(e)}(x;\lambda) &= \sum_{i=1}^{n} c_i^{(j)}(\lambda)\,[a_i(x) + \lambda\,A_{\varepsilon i}(x;\lambda)] \\ &= \sum_{i=1}^{n} c_i^{(j)}(\lambda)\left[a_i(x) + \lambda \int_a^b R_\varepsilon(x,t;\lambda)\,a_i(u)\,\mathrm{d}u\right], \end{aligned} \tag{2.24}$$

where the superscript (e) signifies that $\phi_j^{(e)}(x;\lambda)$ is an eigenfunction of the kernel $K(x,t)$. The span of the set of all eigenfunctions corresponding to the eigenvalue λ constitutes an eigenspace, whose dimension is $p(\lambda)$. The general solution to the homogeneous integral equation corresponding to a given eigenvalue λ assumes the form

$$\phi^{(h)}(x;\lambda) = \sum_{j=1}^{p(\lambda)} \alpha_j \phi_j^{(e)}(x;\lambda),$$

where the α_j are arbitrary constants. The superscript (h) signifies that the solution $\phi^{(h)}(x;\lambda)$ is the most general solution to the homogeneous equation corresponding to the eigenvalue λ.

The number $p(\lambda)$ of linearly independent eigenfunctions corresponding to a given eigenvalue λ is called the *geometric multiplicity of the eigenvalue*. The geometric multiplicity of the eigenvalue is less than or equal to the algebraic multiplicity of the eigenvalue, i.e., if λ is an $m(\lambda)$-fold root of the equation $D_\rho(\lambda) = 0$, then $1 \leq p(\lambda) \leq m(\lambda)$.

When λ is an eigenvalue of a separable kernel, the rationale for considering the homogeneous adjoint integral equation while conducting an analysis of the inhomogeneous integral equation (2.8) was firmly established in Sect. 1.3. Here, we examine this relationship anew when λ is an eigenvalue of a general kernel.

Consider the homogeneous adjoint integral equation

$$\psi(x) = \overline{\lambda} \int_a^b \overline{K(t,x)}\, \psi(t)\, \mathrm{d}t \tag{2.25}$$

where $K(x,t)$ is complex-valued and continuous on the square $Q(a,b)$. If we conjugate the decomposition of the kernel and transpose variables, then we obtain

$$\overline{K(t,x)} = \overline{K_{\mathrm{sep}}(t,x)} + \overline{K_\varepsilon(t,x)}.$$

If we now substitute this decomposition into the homogeneous adjoint equation and rearrange terms, then we have

$$\omega(x) = \psi(x) - \overline{\lambda} \int_a^b \overline{K_\varepsilon(t,x)}\, \psi(t)\, \mathrm{d}t = \overline{\lambda} \int_a^b \overline{K_{\mathrm{sep}}(t,x)}\, \psi(t)\, \mathrm{d}t. \tag{2.26}$$

It is clear from this representation that if $\psi(x) \equiv 0$, then $\omega(x) \equiv 0$ on $[a,b]$. Since the left half of this equation is an inhomogeneous Fredholm equation with a small kernel, the Theorem of Successive Approximation can be applied to it, thereby obtaining a representation for $\psi(x)$ in the form

$$\psi(x) = \omega(x) + \overline{\lambda} \int_a^b \overline{R_\varepsilon(t,x;\lambda)}\, \omega(t)\, \mathrm{d}t, \tag{2.27}$$

where the resolvent kernel $\overline{R_\varepsilon(t,x;\lambda)}$ is the complex conjugate of the resolvent kernel $R_\varepsilon(x,t;\lambda)$ constructed from the iterations of the kernel $K_\varepsilon(x,t)$. It is clear from this representation that if $\omega(x) \equiv 0$, then $\psi(x) \equiv 0$ on $[a,b]$. Thus, $\omega(x)$ and $\psi(x)$ vanish only simultaneously. If we substitute the right half of the representation (2.26) for $\omega(x)$ into the last equation, we obtain

$$\psi(x) = \overline{\lambda} \int_a^b \overline{G_\varepsilon(t,x;\lambda)}\, \psi(t)\, \mathrm{d}t, \tag{2.28}$$

where the kernel $\overline{G_\varepsilon(t,x;\lambda)}$ is the complex conjugate of the kernel $G_\varepsilon(x,t;\lambda)$ that appears in Eq. (2.11). It is explicitly given in the integral form (2.13) or in the equivalent separable form (2.15), from which we conclude that

$$\overline{G_\varepsilon(t,x;\lambda)} = \sum_{i=1}^{n} \overline{[a_i(t) + \lambda A_{\varepsilon i}(t;\lambda)]}\, b_i(x). \tag{2.29}$$

It follows that the solutions to Eq. (2.28) have the representation

$$\psi_j^{(e)}(x;\lambda) = \overline{\lambda} \sum_{i=1}^{n} d_i^{(j)}(\lambda)\, b_i(x),$$

for $j = 1, \ldots, q(\lambda)$, where $q(\lambda)$ is the geometric multiplicity of λ and we have set

$$d_i^{(j)}(\lambda) = \int_a^b \psi(t)\, \overline{[a_i(t) + \lambda A_{\varepsilon i}(t;\lambda)]}\, \mathrm{d}t.$$

We conclude that $\overline{\lambda}$ is an eigenvalue of the kernel $\overline{K(t,x)}$ and that all of its $q(\lambda)$ linearly independent solutions are as given. Since Eq. (2.28) and the equation

$$\phi(x) = \lambda \int_a^b G_\varepsilon(x,t;\lambda)\, \phi(t)\, \mathrm{d}t$$

have conjugate kernels, they must have the same number of linearly independent eigenfunctions, i.e., $p(\lambda) = q(\lambda)$.

The discussion above serves to establish the following result:

Theorem 2.4.3 (The Second Fredholm Theorem). *Let $K(x,t)$ be a complex-valued continuous kernel defined on the square $Q(a,b)$. If λ is an eigenvalue of the kernel $K(x,t)$, then $\overline{\lambda}$ is an eigenvalue of the adjoint kernel $\overline{K(t,x)}$. The number of linearly independent eigenfunctions of the homogeneous equation*

$$\phi(x) = \lambda \int_a^b K(x,t)\,\phi(t)\,\mathrm{d}t$$

is equal to the number of linearly independent eigenfunctions of the homogeneous adjoint equation

$$\psi(x) = \overline{\lambda} \int_a^b \overline{K(t,x)}\,\psi(t)\,\mathrm{d}t.$$

2. If $\mathbf{f}(\lambda) \neq \mathbf{0}$, then neither of the free terms $f_\varepsilon(x;\lambda)$ and $f(x)$ can vanish identically on the interval $[a,b]$, so that the equations

$$\phi(x) = f(x) + \lambda \int_a^b K(x,t)\,\phi(t)\,\mathrm{d}t$$

and

$$\phi(x) = f_\varepsilon(x;\lambda) + \lambda \int_a^b G_\varepsilon(x,t;\lambda)\,\phi(t)\,\mathrm{d}t$$

are simultaneously inhomogeneous and are simultaneously satisfied by $\phi(x)$. Since $G_\varepsilon(x,t;\lambda)$ is separable, the Third Fredholm Theorem for integral equations with separable kernels that was proven in Sect. 1.3 can be invoked to conclude that the last equation has solutions if and only if $f_\varepsilon(x;\lambda)$ is orthogonal to all solutions of the homogeneous adjoint equation

$$\psi(x) = \overline{\lambda} \int_a^b \overline{G_\varepsilon(t,x;\lambda)}\,\psi(t)\,\mathrm{d}t.$$

If $\omega(x)$ denotes one of these solutions, then, given the representation (2.12), we have

$$\begin{aligned}
\int_a^b f_\varepsilon(t;\lambda)\,\overline{\omega(t)}\,\mathrm{d}t &= \int_a^b \left(f(t) + \lambda \int_a^b R_\varepsilon(t,s;\lambda)\,f(s)\,\mathrm{d}s \right) \overline{\omega(t)}\,\mathrm{d}t \\
&= \int_a^b f(t)\,\overline{\omega(t)}\,\mathrm{d}t + \lambda \int_a^b \int_a^b R_\varepsilon(t,s;\lambda)\,f(s)\,\overline{\omega(t)}\,\mathrm{d}s\,\mathrm{d}t \\
&= \int_a^b f(t) \left[\overline{\omega(t)} + \lambda \int_a^b R_\varepsilon(s,t;\lambda)\,\overline{\omega(s)}\,\mathrm{d}s \right] \mathrm{d}t \\
&= \int_a^b f(t) \overline{\left[\omega(t) + \overline{\lambda} \int_a^b \overline{R_\varepsilon(s,t;\lambda)}\,\omega(s)\,\mathrm{d}s \right]}\,\mathrm{d}t \\
&= \int_a^b f(t)\,\overline{\psi(t)}\,\mathrm{d}t.
\end{aligned}$$

The last equality follows from the representation (2.27). We conclude from these equations that $f_\varepsilon(x;\lambda)$ is orthogonal to $\omega(x)$ if and only if $f(x)$ is orthogonal to any solution $\psi(x)$ of the homogeneous adjoint equation

$$\psi(x) = \overline{\lambda} \int_a^b \overline{K(t,x)}\, \psi(t)\, \mathrm{d}t.$$

The discussion above serves to establish the following result:

Theorem 2.4.4 (The Third Fredholm Theorem). *Let λ be a complex parameter, $f(x)$ be a complex-valued continuous function defined on the interval $[a,b]$, and $K(x,t)$ be a complex-valued continuous kernel defined on the square $Q(a,b)$. If λ is an eigenvalue of the kernel $K(x,t)$, then the inhomogeneous Fredholm integral equation of the second kind*

$$\phi(x) = f(x) + \lambda \int_a^b K(x,t)\, \phi(t)\, \mathrm{d}t$$

will have solutions if and only if the free term $f(x)$ is orthogonal to all of the eigenfunctions of the homogeneous adjoint equation

$$\psi(x) = \overline{\lambda} \int_a^b \overline{K(t,x)}\, \psi(t)\, \mathrm{d}t.$$

The Fredholm Alternative Theorem for integral equations with separable kernels was stated in Sect. 1.3. The statement of this theorem for general kernels is identical, except for the fact that the word "separable" is deleted.

Although the Fredholm theorems give the conditions under which solutions to Fredholm integral equations of the second kind exist, they do not specify how to construct those solutions. In the next section, we provide a recursive procedure for constructing the resolvent kernel when λ is regular.

Illustrative Examples

- *Example 1: The Lalesco–Picard Equation*. This example highlights the importance of the Fourth Fredholm Theorem.

  Consider the homogeneous integral equation

  $$\phi(x) = \lambda \int_{-b}^{+b} \mathrm{e}^{-|x-t|}\, \phi(t)\, \mathrm{d}t,$$

  where $0 < \lambda < +\infty$ and $0 < b \leq +\infty$. The presence of the absolute value in the integrand requires the equation to be expanded in the equivalent form

  $$\phi(x) = \lambda \left(\mathrm{e}^{-x} \int_{-b}^{x} \mathrm{e}^{t}\, \phi(t)\, \mathrm{d}t + \mathrm{e}^{x} \int_{x}^{+b} \mathrm{e}^{-t}\, \phi(t)\, \mathrm{d}t \right).$$

  After differentiation and simplification, we obtain the second-order linear ordinary differential equation

$$\phi''(x) + (2\lambda - 1)\,\phi(x) = 0.$$

Any solution to the Lalesco–Picard equation must necessarily be a solution to this differential equation, which has exactly two linearly independent solutions whose form depends upon the value of λ. These solutions are

$$\phi(x) = \begin{cases} A\sinh(\sqrt{1-2\lambda}) + B\cosh(\sqrt{1-2\lambda}), & \text{if } 0 < \lambda < \frac{1}{2} \\ A + Bx, & \text{if } \lambda = \frac{1}{2} \\ A\sin(\sqrt{2\lambda-1}) + B\cos(\sqrt{2\lambda-1}), & \text{if } \frac{1}{2} < \lambda < +\infty \end{cases}$$

It remains to determine, by process of elimination, exactly which of these six solutions to the differential equation are eigenfunctions of the kernel.

We consider three cases:

Case I: $0 < \lambda < \frac{1}{2}$.

In this case, we set $\mu = \sqrt{1-2\lambda}$ or $\lambda = (1-\mu^2)/2$, where $0 < \mu < 1$.

After substituting $\phi(x) = \sinh(\mu x)$ into the expanded integral equation, we obtain

$$\sinh(\mu x) = \sinh(\mu x) + \lambda\,\mathrm{e}^{-b}\sinh x\left(-\frac{\mathrm{e}^{b\mu}}{1-\mu} + \frac{\mathrm{e}^{-b\mu}}{\mu+1}\right);$$

after substituting $\phi(x) = \cosh(\mu x)$, we obtain

$$\cosh(\mu x) = \cosh(\mu x) - \lambda\,\mathrm{e}^{-b}\cosh x\left(\frac{\mathrm{e}^{b\mu}}{1-\mu} + \frac{\mathrm{e}^{-b\mu}}{\mu+1}\right).$$

If $b = +\infty$, then equality holds in both equations, implying that both $\phi_1^{(e)}(x;\lambda) = \sinh(\mu x)$ and $\phi_2^{(e)}(x;\lambda) = \cosh(\mu x)$ are eigenfunctions of the kernel for *every* $\lambda \in (0, \frac{1}{2})$.

However, if $b < +\infty$, then equality cannot hold in either equation for any value of b, so that neither $\sinh(\mu x)$ nor $\cosh(\mu x)$ are eigenfunctions of the kernel for any $\lambda \in (0, \frac{1}{2})$.

Case II: $\lambda = \frac{1}{2}$.

After substituting $\phi(x) = 1$ and then $\phi(x) = x$ into the expanded integral equation, we obtain

$$1 = 1 - \mathrm{e}^{-b} \quad \text{and} \quad x = x - (b+1)\,\mathrm{e}^{-b}\sinh x.$$

If $b = +\infty$, then inequality holds in both equations, implying that both $\phi_1^{(e)}(x;\frac{1}{2}) = 1$ and $\phi_2^{(e)}(x;\frac{1}{2}) = x$ are eigenfunctions of the kernel.

However, if $b < +\infty$, then equality cannot hold in either equation for any value of b, so that neither 1 nor x is an eigenfunction of the kernel.

Case III: $\frac{1}{2} < \lambda < +\infty$.

In this case, we set $\mu = \sqrt{2\lambda - 1}$ or $\lambda = (1+\mu^2)/2$, where $0 < \mu < +\infty$.

After substituting $\phi(x) = \sin(\mu x)$ into the expanded integral equation, we obtain

$$\sin(\mu x) = \sin(\mu x) - \mathrm{e}^{-b} \sinh x \, (\sin(\mu b) + \mu \cos(\mu b));$$

after substituting $\phi(x) = \cos(\mu x)$, we obtain

$$\cos(\mu x) = \cos(\mu x) + \mathrm{e}^{-b} \cosh x \, (\mu \sin(\mu b) - \cos(\mu b)).$$

If $b = +\infty$, then equality holds in both equations implying that both $\phi_1^{(e)}(x;\lambda) = \sin(\mu x)$ and $\phi_2^{(e)}(x;\lambda) = \cos(\mu x)$ are eigenfunctions of the kernel for *every* $\lambda \in (\frac{1}{2}, +\infty)$.

However, if $b < +\infty$, then equality holds in the first equation if and only if $\sin(\mu b) + \mu \cos(\mu b) = 0$, i.e., if $\tan(\mu b) = -\mu$. For each fixed value of b, there are an infinite number of eigenvalues $\lambda_n = (1+\mu_n^2)/2$ corresponding to the eigenfunctions $\phi_1^{(e)}(x;\lambda_n) = \sin(\mu_n x)$. Equality holds in the second equation if and only if $\mu \sin(\mu b) - \cos(\mu b) = 0$, i.e., if $\tan(\mu b) = 1/\mu$. For each fixed value of b, there are an infinite number of eigenvalues $\lambda_n = (1+\mu_n^2)/2$ corresponding to the eigenfunctions $\phi_2^{(e)}(x;\lambda_n) = \cos(\mu_n x)$.

In summary, we have come to the following conclusions:

- If $b = +\infty$, then *every* $\lambda \in (0, +\infty)$ is an eigenvalue of the kernel. Two eigenfunctions correspond to each eigenvalue λ.
- If $b < +\infty$, then there is a countably infinite set of discrete eigenvalues $\{\lambda_n\}$ which do not have a finite limit point, in perfect agreement with the Fourth Fredholm Theorem. One eigenfunction corresponds to each eigenvalue λ_n.

2.5 Constructing the Resolvent Kernel

In this section, we prove the following result:

Theorem 2.5.1. *The resolvent kernel $R(x,t;\lambda)$ can be written as the quotient of the entire functions $D(x,t;\lambda)$ and $D(\lambda)$ given by the series expansions*

$$D(x,t;\lambda) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} B_n(x,t) \lambda^n \quad \text{and} \quad D(\lambda) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} c_n \lambda^n.$$

In these series, $B_0(x,t) = K(x,t)$ and $c_0 = 1$. For every $n \geq 1$, $B_n(x,t)$ can be computed from the recursive relationship

$$B_n(x,t) = c_n K(x,t) - n \int_a^b K(x,s) B_{n-1}(s,t) \, \mathrm{d}s,$$

and c_n can be computed by evaluating the integral

$$c_n = \int_a^b B_{n-1}(t,t)\,\mathrm{d}t.$$

Before beginning the proof, we remark that the coefficients $B_n(x,t)$ and c_n in these series are computed sequentially. Schematically, the sequence of successive integrations are represented by the diagram

$$\begin{array}{ccccccc}
B_0 & \to & c_1 & & & & \\
\downarrow & \swarrow & & & & & \\
B_1 & \to & c_2 & & & & \\
\downarrow & \swarrow & & & & & \\
B_2 & \to & c_3 & & & & \\
\downarrow & \swarrow & & & & & \\
B_3 & \to & c_4 & & & & \\
\downarrow & \swarrow & & & & & \\
B_4 & \to & c_5 & & & & \\
\vdots & & \vdots & & \vdots & &
\end{array}$$

Note that if $B_N(x,t)$ vanishes identically for some N, then $c_{N+1} = 0$, and $B_n(x,t)$ also vanishes identically for all $n > N$. In this case, both series are just polynomials in λ of degree $N-1$ and N, respectively.

Proof. In order to show that the resolvent $R(x,t;\lambda)$ can be expressed as the quotient of two specific entire functions of λ, we require the following technical lemma which measures the difference between the resolvent kernels $R(x,t;\lambda)$ generated by iterating $K(x,t)$ and $R_\varepsilon(x,t;\lambda)$ generated by iterating $K_\varepsilon(x,t)$.

Lemma 2.5.1. *For all values of $\lambda \in \Delta_\rho$, we have*

$$\int_a^b U_\varepsilon(t,t;\lambda)\,\mathrm{d}t = \int_a^b R_\varepsilon(t,t;\lambda)\,\mathrm{d}t - \frac{D_\rho{}'(\lambda)}{D_\rho(\lambda)}.$$

Proof. The proof consists of a sequence of equalities requiring Eq. (2.22) in addition to the definitions of $S_\varepsilon(x,t;\lambda)$, $a_{ji}(\lambda)$, $A_{\varepsilon i}(x;\lambda)$, $a'_{ji}(\lambda)$, and the integrodifferential equation for the resolvent (see Sect. 2.3, Exercise 9). We have

$$\begin{aligned}
&\int_a^b U_\varepsilon(t,t;\lambda)\,\mathrm{d}t - \int_a^b R_\varepsilon(t,t;\lambda)\,\mathrm{d}t \\
&\quad = \int_a^b S_\varepsilon(t,t;\lambda) + \lambda \int_a^b \int_a^b S_\varepsilon(t,s;\lambda)\,R_\varepsilon(s,t;\lambda)\,\mathrm{d}s\,\mathrm{d}t \\
&\quad = \frac{1}{D_\rho(\lambda)} \sum_{i=1}^n \sum_{j=1}^n D_{ji}(\lambda) \int_a^b \left[a_i(t) + \lambda A_{\varepsilon i}(t;\lambda)\right] b_j(t)\,\mathrm{d}t
\end{aligned}$$

$$
\begin{aligned}
&+\lambda \sum_{i=1}^{n}\sum_{j=1}^{n} D_{ji}(\lambda)\int_a^b\int_a^b [a_i(t)+\lambda A_{\varepsilon i}(t;\lambda)]\,b_j(s)R_\varepsilon(s,t;\lambda)\,\mathrm{d}s\,\mathrm{d}t\\
&=\frac{1}{D_\rho(\lambda)}\sum_{i=1}^{n}\sum_{j=1}^{n} D_{ji}(\lambda)a_{ji}(\lambda)\\
&\quad+\lambda\sum_{i=1}^{n}\sum_{j=1}^{n} D_{ji}(\lambda)\left[\int_a^b\int_a^b R_\varepsilon(s,t;\lambda)a_i(t)b_j(s)\,\mathrm{d}t\,\mathrm{d}s\right.\\
&\quad\left.+\int_a^b\int_a^b\left(\int_a^b R_\varepsilon(s,t;\lambda)R_\varepsilon(t,u;\lambda)\,\mathrm{d}t\right)a_i(u)b_j(s)\,\mathrm{d}s\,\mathrm{d}u\right]\\
&=\frac{1}{D_\rho(\lambda)}\sum_{i=1}^{n}\sum_{j=1}^{n} D_{ji}(\lambda)a_{ji}(\lambda)+\lambda\sum_{i=1}^{n}\sum_{j=1}^{n} D_{ji}(\lambda)\left[\int_a^b A_{\varepsilon i}(s;\lambda)b_j(s)\right.\\
&\quad\left.+\lambda\int_a^b\int_a^b\frac{\partial}{\partial\lambda}R_\varepsilon(s,u;\lambda)a_i(u)b_j(s)\,\mathrm{d}u\,\mathrm{d}s\right]\\
&=\frac{1}{D_\rho(\lambda)}\sum_{i=1}^{n}\sum_{j=1}^{n} D_{ji}(\lambda)a_{ji}(\lambda)\\
&\quad+\lambda\frac{1}{D_\rho(\lambda)}\sum_{i=1}^{n}\sum_{j=1}^{n} D_{ji}(\lambda)\\
&\quad\times\int_a^b\left(A_{\varepsilon i}(s;\lambda)+\lambda\frac{\partial}{\partial\lambda}\left(\int_a^b R_\varepsilon(s,u;\lambda)a_i(u)\,\mathrm{d}u\right)\right)b_j(s)\,\mathrm{d}s\\
&=\frac{1}{D_\rho(\lambda)}\left[\sum_{i=1}^{n}\sum_{j=1}^{n} D_{ji}(\lambda)\left(a_{ji}(\lambda)+\lambda a'_{ji}(\lambda)\right)\right]\\
&=-\frac{1}{D_\rho(\lambda)}\sum_{i=1}^{n}\sum_{j=1}^{n} D_{ji}(\lambda)\left(-\lambda a_{ji}(\lambda)\right)'\\
&=-\frac{D_\rho{}'(\lambda)}{D_\rho(\lambda)}.
\end{aligned}
$$

□

This lemma will be used to show that there exist entire functions $D(x,t;\lambda)$ and $D(\lambda)$ such that

$$D(x,t;\lambda)=R(x,t;\lambda)D(\lambda).$$

Since $R(t,t;\lambda)$ is an analytic function of λ in Δ_ρ, the statement of the lemma implies that the zeroes of $D_\rho(\lambda)$ correspond to the poles of the function

$$\delta_\varepsilon(\lambda)=\int_a^b U_\varepsilon(t,t;\lambda)\,\mathrm{d}t$$

that lie in Δ_ρ. Specifically, if $\lambda_k \in \Delta_\rho$ is a zero of algebraic multiplicity m_k in $D_\rho(\lambda)$, then $D_\rho(\lambda) = (\lambda - \lambda_k)^{m_k} e_\rho(\lambda)$, where $e_\rho(\lambda) \neq 0$ near λ_k, and

$$\frac{D_\rho{}'(\lambda)}{D_\rho(\lambda)} = \frac{m_k}{\lambda - \lambda_k} + \frac{e_\rho{}'(\lambda)}{e_\rho(\lambda)}.$$

This fact in conjunction with the equality stated in the lemma shows that λ_k is a simple pole of $\delta_\varepsilon(\lambda)$ with residue $-m_k$. Hence, $\delta_\varepsilon(\lambda)$ has only simple poles which coincide with the zeroes of $D_\rho(\lambda)$. Since $U_\varepsilon(t,t;\lambda)$ is the restriction of $R(t,t;\lambda)$ to the disk Δ_ρ, $\delta_\varepsilon(\lambda)$ is also the restriction of

$$\delta(\lambda) = \int_a^b R(t,t;\lambda)\,\mathrm{d}t$$

to the disk Δ_ρ, so that $\delta(\lambda)$ is a meromorphic function in the full λ-plane having only simple poles that coincide with the eigenvalues of the kernel $K(x,t)$. From this, it follows that

$$D(\lambda) = \exp\left(-\int_0^\lambda \delta(\lambda)\,\mathrm{d}\lambda\right)$$

is an entire function of λ.

To show that $D(x,t;\lambda)$ is entire, note that we can integrate the equality stated in the lemma near the origin since $D_\rho(0) = 1$ to obtain the relation

$$-\int_0^\lambda \delta(\lambda)\,\mathrm{d}\lambda = -\int_0^\lambda \int_a^b R_\varepsilon(t,t;\lambda)\,\mathrm{d}t\,\mathrm{d}\lambda + \ln D_\rho(\lambda)$$

which is valid in Δ_ρ. Upon exponentiation, we obtain

$$D(\lambda) = D_\rho(\lambda) \cdot \exp\left(-\int_0^\lambda \int_a^b R_\varepsilon(t,t;\lambda)\,\mathrm{d}t\,\mathrm{d}\lambda\right).$$

It follows from this relation that the zeroes of $D(\lambda)$ coincide with those of $D_\rho(\lambda)$ in Δ_ρ and have the same multiplicity, since the exponential factor does not vanish. Thus, $D(x,t;\lambda)$ is an entire function of λ, since the poles of $R(t,t;\lambda)$ are cancelled by the zeroes of $D(\lambda)$.

It remains to develop the series expansions in the statement of the theorem. Recall from Sect. 2.3 that the resolvent kernel can be represented by the expansion

$$R(x,t;\lambda) = \sum_{m=1}^{\infty} \lambda^{m-1} K_m(x,t)$$

for small values of λ, where $K_m(x,t)$ is the m-th iterated kernel of $K(x,t)$. It follows that

$$\delta(\lambda) = \int_a^b R(t,t;\lambda)\,\mathrm{d}t = \sum_{m=1}^{\infty} A_m \lambda^{m-1},$$

where we have set

$$A_m = \int_a^b K_m(t,t)\,\mathrm{d}t.$$

The coefficient A_m is known as the *trace of the kernel* $K_m(x,t)$. It follows that

$$\begin{aligned}
D(\lambda) &= \exp\left[-\int_0^\lambda \left(\sum_{m=1}^{\infty} A_m \lambda^{m-1}\right)\mathrm{d}\lambda\right] \\
&= \exp\left(-\sum_{m=1}^{\infty}\frac{A_m}{m}\lambda^m\right) \\
&= \sum_{k=0}^{\infty}\frac{(-1)^k}{k!}\left(\sum_{m=1}^{\infty}\frac{A_m}{m}\lambda^m\right) \\
&= 1+\frac{(-1)^1}{1!}A_1\lambda+\frac{(-1)^2}{2!}\left({A_1}^2-A_2\right)\lambda^2 \\
&\quad+\frac{(-1)^3}{3!}\left({A_1}^3-3A_1A_2+2A_3\right)\lambda^3 \\
&\quad+\frac{(-1)^4}{4!}\left({A_1}^4-6{A_1}^2A_2+3{A_2}^2+8A_1A_3-6A_4\right)\lambda^4+\cdots.
\end{aligned}$$

More simply, we can write

$$D(\lambda) = \sum_{n=0}^{\infty}\frac{(-1)^n}{n!}c_n\lambda^n,$$

where $c_0 = 1$. Also, for $n \geq 2$, each coefficient c_n is a evidently a multivariable polynomial in the traces $A_1,\ldots,A_n$ whose coefficients are integers that sum to zero. (To establish this purely algebraic statement, just set $A_m = 1$ for $m \geq 1$. Then $D(\lambda) = 1-\lambda$, so that $c_n = 0$ for all $n \geq 2$.)

Since $D(x,t;\lambda)$ is entire, it has the Maclaurin expansion

$$D(x,t;\lambda) = \sum_{n=0}^{\infty}\frac{1}{n!}D^{(n)}(x,t;0)\lambda^n = \sum_{n=0}^{\infty}\frac{(-1)^n}{n!}B_n(x,t)\lambda^n.$$

The coefficients $B_n(x,t)$ in this series must be computed in order to establish the recursion relation for them given in the statement of the theorem. If $n = 0$, then

$$B_0(x,t) = D(x,t;0) = R(x,t;0)D(0) = K(x,t).$$

If $n \geq 1$, then we apply Leibniz formula to obtain

$$\begin{aligned}
B_n(x,t) &= (-1)^n D^{(n)}(x,t;0) \\
&= (-1)^n\left(D(0)R(x,t;0)\right)^{(n)}
\end{aligned}$$

$$
\begin{aligned}
&= (-1)^n \sum_{m=0}^{n} \frac{n!}{m!\,(n-m)!}\, D^{(n-m)}(0)\, R^{(m)}(x,t;0) \\
&= \sum_{m=0}^{n} \frac{n!}{m!\,(n-m)!}\, \left((-1)^{-m} c_{n-m}\right) \left(m!\, K_{m+1}(x,t)\right) \\
&= n! \sum_{m=0}^{n} (-1)^{-m} \frac{1}{(n-m)!}\, c_{n-m}\, K_{m+1}(x,t) \\
&= n! \left(\frac{1}{n!}\, c_n K(x,t) + \sum_{m=1}^{n} \frac{(-1)^{-m}}{(n-m)!}\, c_{n-m} K_{m+1}(x,t) \right) \\
&= c_n K(x,t) - n \left((n-1)! \sum_{m=0}^{n-1} \frac{(-1)^{-m}}{((n-1)-m)!}\, c_{(n-1)-m} K_{m+2}(x,t) \right) \\
&= c_n K(x,t) \\
&\quad - n \left[\int_a^b K(x,s) \right. \\
&\quad \left. \times \left((n-1)! \sum_{m=0}^{n-1} \frac{(-1)^{-m}}{((n-1)-m)!}\, c_{(n-1)-m} K_{m+1}(s,t) \right) \mathrm{d}s \right] \\
&= c_n K(x,t) - n \int_a^b K(x,s)\, B_{n-1}(s,t)\, \mathrm{d}s.
\end{aligned}
$$

This sequence of equalities shows that $B_n(x,t)$ is a particular linear combination of iterated kernels and that $B_n(x,t)$ can be computed recursively.

Finally, we show that each c_{n+1} may be determined from the coefficient function $B_n(x,t)$. On the one hand, we have

$$
\begin{aligned}
\int_a^b D(t,t;\lambda)\, \mathrm{d}t &= D(\lambda) \int_a^b R(t,t;\lambda)\, \mathrm{d}t \\
&= D(\lambda)\, \delta(\lambda) \\
&= -D'(\lambda) \\
&= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!}\, c_{n+1}\, \lambda^n.
\end{aligned}
$$

On the other hand, we also have

$$
\int_a^b D(t,t;\lambda)\, \mathrm{d}t = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \left(\int_a^b B_n(t,t)\, \mathrm{d}t \right) \lambda^n.
$$

By comparing the coefficients of the powers of λ in these two expansions, we have

$$c_{n+1} = \int_a^b B_n(t,t)\,\mathrm{d}t$$

for all $n \geq 0$ as required. □

The coefficients c_n are related to the traces A_m in the following way. By setting $x = t$ in the sequence of equalities in the proof, we obtain the relation

$$B_n(t,t) = n!\left(\sum_{m=0}^{n}(-1)^m \frac{1}{(n-m)!}\,c_{n-m}\,K_{m+1}(t,t)\right).$$

Upon integration, we immediately have

$$c_{n+1} = n!\left(\sum_{m=0}^{n}(-1)^m \frac{1}{(n-m)!}\,c_{n-m}\,A_{m+1}\right).$$

It is of interest to note that the coefficients c_n can be written as integrals of determinants. By definition, we have

$$c_1 = A_1 = \int_a^b K(t,t)\,\mathrm{d}t.$$

Similarly, we have

$$\begin{aligned} c_2 &= {A_1}^2 - A_2 \\ &= \left(\int_a^b K(t_1,t_1)\,\mathrm{d}t_1\right)^2 - \int_a^b K_2(t_1,t_1)\,\mathrm{d}t_1 \\ &= \left(\int_a^b K(t_1,t_1)\,\mathrm{d}t_1\right)\left(\int_a^b K(t_2,t_2)\,\mathrm{d}t_2\right) - \int_a^b\int_a^b K(t_1,t_2)\,K(t_2,t_1)\,\mathrm{d}t_2\,\mathrm{d}t_1 \\ &= \int_a^b\int_a^b \det\begin{pmatrix} K(t_1,t_1) & K(t_1,t_2) \\ K(t_2,t_1) & K(t_2,t_2) \end{pmatrix}\mathrm{d}t_2\,\mathrm{d}t_1. \end{aligned}$$

In general, it can be shown that

$$c_n = \int_a^b \cdots \int_a^b \det\begin{pmatrix} K(t_1,t_1) & K(t_1,t_2) & \cdots & K(t_1,t_n) \\ K(t_2,t_1) & K(t_2,t_2) & \cdots & K(t_1,t_n) \\ \vdots & \vdots & \ddots & \vdots \\ K(t_n,t_1) & K(t_n,t_2) & \cdots & K(t_n,t_n) \end{pmatrix}\mathrm{d}t_n \ldots \mathrm{d}t_1.$$

The magnitude of the coefficients $|c_n|$ can be estimated with a classical inequality due to Hadamard stating that the absolute value of the determinant of a matrix is less than or equal to the product of the norms of its rows. More precisely, if $\mathbf{A} = (a_{ij})$ is an $n \times n$ matrix and $\mathbf{a}_i = (a_{i1}, \ldots, a_{in})$ is its ith row, then

$$|\det \mathbf{A}| \leq \prod_{i=1}^{n} \|\mathbf{a}_i\|,$$

where $\|\mathbf{a}_i\| = (a_{i1}^2 + \cdots + a_{in}^2)^{1/2}$. By Hadamard's inequality, we have

$$|c_n| \leq n^{n/2} M^n (b-a)^n,$$

if $|K(x,t)| \leq M$ on the square $Q(a,b)$.

The coefficient functions $B_n(x,t)$ can also be written as multiple integrals of a determinant. Specifically, for $n \geq 1$, we have

$$B_n(x,t) = \int_a^b \ldots \int_a^b \Delta_n(x,t)\, \mathrm{d}t_n, \ldots, \mathrm{d}t_1,$$

where

$$\Delta_n(x,t) = \det \begin{pmatrix} K(x,t) & K(x,t_1) & K(x,t_2) & \ldots & K(x,t_n) \\ K(t_1,t) & K(t_1,t_1) & K(t_1,t_2) & \ldots & K(t_1,t_n) \\ K(t_2,t) & K(t_2,t_1) & K(t_2,t_2) & \ldots & K(t_1,t_n) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ K(t_n,t) & K(t_n,t_1) & K(t_n,t_2) & \ldots & K(t_n,t_n) \end{pmatrix}.$$

Illustrative Examples

- *Example 1*: Recall the inhomogeneous Fredholm integral equation of the second kind

$$\phi(x) = f(x) + \lambda \int_0^1 (xt^2 + x^2 t^4)\, \phi(t)\, \mathrm{d}t$$

that was solved in Sect. 1.2. It is always the case that $c_0 = 1$. Since $B_0(x,t) = K(x,t) = xt^2 + x^2 t^4$, we have

$$c_1 = \int_0^1 B_0(t,t)\, \mathrm{d}t = \int_0^1 (t^3 + t^6)\, \mathrm{d}t = \tfrac{11}{28}.$$

Also,

$$\begin{aligned} B_1(x,t) &= c_1 K(x,t) - 1 \cdot \int_0^1 K(x,s) B_0(s,t)\,\mathrm{d}t \\ &= \tfrac{11}{28}(xt^2 + x^2t^4) - \int_0^1 (xs^2 + x^2s^4)(st^2 + s^2t^4)\,\mathrm{d}t \\ &= \tfrac{1}{4}x^2t^4 + \tfrac{1}{7}xt^2 - \tfrac{1}{5}xt^4 - \tfrac{1}{6}x^2t^2 \end{aligned}$$

and

$$c_2 = \int_0^1 B_1(t,t)\,\mathrm{d}t = \int_0^1 (\tfrac{1}{4}t^6 + \tfrac{1}{7}t^3 - \tfrac{1}{5}t^5 - \tfrac{1}{6}t^4)\,\mathrm{d}t = \tfrac{1}{210}.$$

One more integration shows that $B_2(x,t) \equiv 0$, so that $c_n = 0$ and $B_n(x,t) \equiv 0$ if $n \geq 3$. Hence,

$$D(\lambda) = 1 + (-1)\frac{1}{1!}\left(\tfrac{11}{28}\right)\lambda + (-1)^2\frac{1}{2!}\left(\tfrac{1}{210}\right)\lambda^2 = \tfrac{1}{420}\left(\lambda^2 - 165\lambda + 420\right)$$

and

$$\begin{aligned} D(x,t;\lambda) &= B_0(x,t) + (-1)\frac{1}{1!}B_1(x,t)\lambda \\ &= (xt^2 + x^2t^4) - \left(\tfrac{1}{7}xt^2 + \tfrac{1}{4}x^2t^4 - \tfrac{1}{5}xt^4 - \tfrac{1}{6}x^2t^2\right)\lambda. \end{aligned}$$

These results are identical to those obtained in Sect. 1.2.

Section 2.5 Exercises

1. Let $K(x,t) = a(x)\overline{b(t)}$, and assume that $\alpha = \int_a^b a(t)\overline{b(t)}\,\mathrm{d}t \neq 0$. Show that

$$R(x,t;\lambda) = \frac{K(x,t)}{1 - \lambda\alpha}$$

and that the mth trace $A_m = \alpha^m$. In particular, if $b(t) = a(t)$, then $A_m = \|a\|_2^{2m}$.

2. Let $K(x,t) = a_1(x)\overline{b_1(t)} + a_2(x)\overline{b_2(t)}$ and define the matrix $\mathbf{A} = (a_{ij})$ by

$$a_{ij} = \int_a^b a_j(t)\overline{b_i(t)}\,\mathrm{d}t.$$

Determine the resolvent kernel $R(x,t;\lambda)$ and show that

$$\int_a^b R(t,t;\lambda)\,\mathrm{d}t = \frac{\mathrm{tr}(\mathbf{A}) - 2\det(\mathbf{A})\lambda}{1 - \mathrm{tr}(\mathbf{A})\lambda + \det(\mathbf{A})\lambda^2} = \sum_{m=1}^{\infty} A_m \lambda^{m-1}$$

if $|\lambda| < \min\{|\lambda_1|, |\lambda_2|\}$. Thus, the mth trace A_m can be expressed in terms of $\mathrm{tr}(\mathbf{A})$ and $\det(\mathbf{A})$.

Also, since

$$D(\lambda) = 1 - \mathrm{tr}(\mathbf{A})\,\lambda + \det(\mathbf{A})\,\lambda^2 = \left(1 - \frac{\lambda}{\lambda_1}\right)\left(1 - \frac{\lambda}{\lambda_2}\right),$$

we have

$$\mathrm{tr}(\mathbf{A}) = \frac{1}{\lambda_1} + \frac{1}{\lambda_2} \quad \text{and} \quad \det(\mathbf{A}) = \frac{1}{\lambda_1\,\lambda_2}.$$

Use this observation to show that

$$A_m = \frac{1}{\lambda_1^m} + \frac{1}{\lambda_2^m}$$

for every $m \geq 1$.

2.6 Numerical Methods

We have seen that considerable difficulties may arise in the computation of the resolvent kernel, even with relatively simple continuous general kernels. Thus, the use of numerical methods assumes critical importance in the production of approximate solutions to Fredholm integral equations of the second kind. In this section, we consider several elementary approaches to the problem of determining approximations to the solution of an integral equation.

2.6.1 The Method of Uniform Approximation

The essence of the theorem in this subsection is that if the two free terms and the two kernels in two Fredholm integral equations are close, then their solutions are close as well. This theorem is extremely useful if one of the integral equations is difficult to solve while the other one is not.

Theorem 2.6.1 (Theorem of Uniform Approximation). *Let $f(x)$ and $g(x)$ be continuous free terms defined on the interval $[a,b]$, and let $K(x,t)$ and $L(x,t)$ be continuous kernels defined on the square $Q(a,b)$.*

Suppose that

$$\max_{a\leq x\leq b} |f(x) - g(x)| < \varepsilon \quad \textit{and} \quad \max_{a\leq x,t\leq b} |K(x,t) - L(x,t)| < \kappa$$

for some fixed constants ε and κ. Suppose further that λ is a common regular value of both kernels of the Fredholm integral equations

$$\phi(x) = f(x) + \lambda \int_a^b K(x,t)\,\phi(t)\,\mathrm{d}t$$

and

$$\psi(x) = g(x) + \lambda \int_a^b L(x,t)\,\psi(t)\,\mathrm{d}t$$

for which $|\lambda|$ *is less than the radii of convergence of both of the resolvent series* $R_K(x,t;\lambda)$ *and* $R_L(x,t;\lambda)$.

Then there exist constants $\alpha = \alpha(\lambda, B_L)$ *and* $\beta = \beta(\lambda, B_f, B_K, B_L)$ *such that*

$$\max_{a\le x\le b} |\phi(x) - \psi(x)| \le \alpha\,\varepsilon + \beta\,\kappa,$$

where

$$B_f = \max_{a\le x\le b} |f(x)|,$$

$$B_K(\lambda) = \max_{a\le x\le b} \left| \int_a^b R_K(x,t;\lambda)\,\mathrm{d}t \right|,$$

and

$$B_L(\lambda) = \max_{a\le x\le b} \left| \int_a^b R_L(x,t;\lambda)\,\mathrm{d}t \right|.$$

Proof. By virtue of the First Fredholm Theorem in Sect. 2.4, the solution to the first equation in the statement of the theorem has the representation

$$\phi(x) = f(x) + \lambda \int_a^b R_K(x,t;\lambda)\,f(t)\,\mathrm{d}t.$$

It follows from this representation that

$$B_\phi = \max_{a\le x\le b} |\phi(x)| \le B_f\,(1 + |\lambda|\,B_K(\lambda)).$$

Next, if we subtract the two integral equations given in the statement of the theorem, we obtain

$$\phi(x) - \psi(x) = D(x) + \lambda \int_a^b L(x,t)\,(\phi(t) - \psi(t))\,\mathrm{d}t,$$

where

$$D(x) = f(x) - g(x) + \lambda \int_a^b (K(x,t) - L(x,t))\,\phi(t)\,\mathrm{d}t.$$

It will be useful to have the estimate

$$B_D = \max_{a\le x\le b} |D(x)| \le \varepsilon + |\lambda|\,\kappa\,(b-a)\,B_\phi.$$

Again, by the First Fredholm Theorem, the solution to the above equation has the representation

$$\phi(x) - \psi(x) = D(x) + \lambda \int_a^b R_L(x,t;\lambda)\, D(t)\, \mathrm{d}t.$$

It follows from this representation that

$$\max_{a \le x \le b} |\phi(x) - \psi(x)| \le (1 + |\lambda|\, B_L(\lambda))\, B_D.$$

Upon combining the above estimates, we finally obtain

$$\max_{a \le x \le b} |\phi(x) - \psi(x)| \le \alpha\, \varepsilon + \beta\, \kappa$$

where

$$\alpha = 1 + |\lambda|\, B_L(\lambda)$$

and

$$\beta = |\lambda|\, (1 + |\lambda|\, B_L(\lambda))\, (1 + |\lambda|\, B_K(\lambda))\, (b-a) B_f. \qquad \square$$

In order to use this theorem, it is necessary to obtain estimates for $B_K(\lambda)$ and $B_L(\lambda)$. If $K(x,t) \le M$ on the square $Q(a,b)$ and λ is within the radius of convergence of the resolvent series, then

$$\left| \int_a^b R_K(x,t;\lambda)\, \mathrm{d}t \right| \le \frac{M(b-a)}{1 - |\lambda| M(b-a)} = B_K(\lambda).$$

Illustrative Examples

- *Example 1*: Consider the integral equation

$$\phi(x) = \cos x + \tfrac{1}{2} \int_0^1 \cos(xt)\, \phi(t)\, \mathrm{d}t.$$

The kernel $K(x,t) = \cos(xt)$ is not separable, and it is difficult to compute its iterated kernels. Consequently, an approximative method is indicated.

Since the free term $\cos x$ can be represented with the Taylor series

$$\cos x = 1 - \frac{1}{2!} x^2 + \frac{1}{4!} x^4 - \frac{1}{6!} x^6 + \cdots,$$

we choose

$$g(x) = 1 - \frac{1}{2!} x^2 + \frac{1}{4!} x^4 \quad \text{and} \quad L(x,t) = 1 - \frac{1}{2!} x^2 t^2 + \frac{1}{4!} x^4 t^4$$

and consider the companion equation

$$\psi(x) = 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 + \frac{1}{2}\int_0^1 \left(1 - \frac{1}{2!}x^2t^2 + \frac{1}{4!}x^4t^4\right)\psi(t)\,\mathrm{d}t.$$

The companion equation has a separable kernel. It can be solved by employing the method prescribed in Sect. 1.3 to obtain

$$\psi(x) = \frac{11,532,090}{6,397,711} - \frac{7,944,195}{12,795,422}x^2 + \frac{607,005}{12,795,422}x^4.$$

If we substitute this approximation into the equation, then we obtain

$$\begin{aligned}
&\psi(x) - \cos x - \frac{1}{2}\int_0^1 \cos(xt)\,\psi(t)\,\mathrm{d}t \\
&\quad = \frac{1,620,362,251}{1,064,067,293,520}x^6 - \frac{1,473,509,027}{55,331,499,263,040}x^8 + O(x^9) \\
&\quad \approx 0.0015228x^6 - 0.0000266x^8
\end{aligned}$$

as the residual. The constant term and the coefficients of x^2 and x^4 in this difference vanish.

The actual error is well within the error predicted by the theorem. Since the Taylor series for $\cos x$ is alternating, we have the classical estimates

$$|f(x) - g(x)| = \left|\cos x - \left(1 - \frac{1}{2}x^2 + \frac{1}{24}x^4\right)\right| < \varepsilon = \frac{1}{720}$$

on the interval $[0,1]$ and

$$|K(x,t) - L(x,t)| = \left|\cos(xt) - \left(1 - \frac{1}{2}x^2t^2 + \frac{1}{24}x^4t^4\right)\right| < \kappa = \frac{1}{720}$$

on the square $Q(0,1)$. Also, $B_f = 1$ and $B_K(\frac{1}{2}) = B_L(\frac{1}{2}) = 2$. Thus, the difference between the actual and the approximate solutions is predicted by the theorem to be no more than $|\phi(x) - \psi(x)| \leq \frac{1}{180} = 0.0055555$.

2.6.2 The Method of Collocation

The Method of Collocation produces a function $y^*(x)$ that approximates the solution $\phi(x)$ to a Fredholm integral equation on an interval as a combination of a predetermined set of continuous and linearly independent functions on that interval.

Although this method is described here for Fredholm integral equations, we note that it can be adapted to produce an approximation to the solution of other types of integral equations as well.

Theorem 2.6.2 (The Collocation Theorem). *Consider the integral equation*

$$\phi(x) = f(x) + \lambda \int_a^b K(x,t)\,\phi(t)\,\mathrm{d}t,$$

where the free term $f(x)$ is continuous on the interval $[a,b]$, the kernel $K(x,t)$ is continuous on the square $Q(a,b)$, and λ is a regular value of the kernel that is less than the radius of convergence of the resolvent series.

Choose a selection $\{x_1,\ldots,x_n\}$ of n nodes with $a \le x_1 < \cdots < x_n \le b$ and a set $\{y_1(x),\ldots,y_n(x)\}$ of n functions that are continuous and linearly independent on the interval $[a,b]$. Define the matrices $\mathbf{Y} = (y_{ik})$ and $\mathbf{Z} = (z_{ik})$ where

$$y_{ik} = y_k(x_i) \quad \textit{and} \quad z_{ik} = \int_a^b K(x_i,t)\,y_k(t)\,\mathrm{d}t.$$

If $\det(\mathbf{Y} - \lambda\,\mathbf{Z}) \neq 0$*, then there exists a linear combination*

$$y^*(x) = \sum_{i=1}^{n} a_i^*\,y_i(x)$$

for which

$$y^*(x_i) = f(x_i) + \lambda \int_a^b K(x_i,t)\,y^*(t)\,\mathrm{d}t$$

for all $i = 1,\ldots,n$. Furthermore, the inequality

$$|\phi(x) - y^*(x)| \le (1 + |\lambda|\,B_R(\lambda))\,B_{y^*}$$

holds uniformly on the interval $[a,b]$, where

$$B_R(\lambda) = \max_{a\le x\le b} \left| \int_a^b R_K(x,t;\lambda)\,\mathrm{d}t \right|$$

and

$$B_{y^*} = \max_{a\le x\le b} \left| y^*(x) - f(x) - \lambda \int_a^b K(x,t)\,y^*(t)\,\mathrm{d}t \right|.$$

Proof. For any choice of $y(x)$, define the *residual* $\eta(x)$ *of* $y(x)$ to be

$$\eta(x) = y(x) - f(x) - \lambda \int_a^b K(x,t)\,y(t)\,\mathrm{d}t.$$

The residual $\eta(x) \equiv 0$ on $[a,b]$ if and only if $y(x)$ is equal to the unique solution $\phi(x)$ of the integral equation.

Let $\{a_1,\ldots,a_n\}$ be an arbitrary choice of coefficents and consider the linear combination

$$y(x) = y(x;a_1,\ldots,a_n) = \sum_{k=1}^{n} a_k y_k(x).$$

For this choice of coefficients, the residual of $y(x)$ assumes the corresponding form

$$\eta(x) = \eta(x;a_1,\ldots,a_n) = \sum_{k=1}^{n} a_k \left(y_k(x) - \lambda z_k(x)\right) - f(x).$$

If $y(x)$ is not the solution to the integral equation, then the residual $\eta(x)$ does not identically vanish on the interval $[a,b]$. However, if $\eta(x_i) = 0$ for all $i = 1,\ldots,n$, then the approximate solution $y(x)$ will be measureably close to the actual solution $\phi(x)$ on the entire interval $[a,b]$. It is this idea which constitutes the essence of the Method of Collocation.

By requiring that $\eta(x_i) = 0$ for all $i = 1,\ldots,n$, we produce a system of n linear equations in the n unknowns $a_1,\ldots,a_n$. This linear system may be written in matrix form as

$$(\mathbf{Y} - \lambda\,\mathbf{Z})\,\mathbf{a} = \mathbf{f}$$

where $\mathbf{a} = (a_1,\ldots,a_n)^{\mathrm{T}}$ and $\mathbf{f} = (f(x_1),\ldots,f(x_n))^{\mathrm{T}}$. If $\det(\mathbf{Y} - \lambda\,\mathbf{Z}) \neq 0$, then the linear system has a unique solution

$$\mathbf{a}^* = (a_1^*,\ldots,a_n^*)^{\mathrm{T}} = (\mathbf{Y} - \lambda\,\mathbf{Z})^{-1}\,\mathbf{f}.$$

The corresponding approximate solution

$$y^*(x) = y^*(x;a_1^*,\ldots,a_n^*) = \sum_{k=1}^{n} a_k^* y_k(x)$$

has the property that its residual vanishes at the selected nodes x_i, i.e., that $\eta^*(x_i) = \eta^*(x_i;a_1^*,\ldots,a_n^*) = 0$ for all $i = 1,\ldots,n$.

For any choice of $y(x)$, let $\varepsilon(x) = y(x) - \phi(x)$. By adding the integral equation to the definition of the residual of $y(x)$, we obtain

$$\varepsilon(x) = \eta(x) + \lambda \int_a^b K(x,t)\,\varepsilon(t)\,\mathrm{d}t.$$

Since $\eta(x)$ continuous on the interval $[a,b]$, this equation has a unique solution that can be represented in the form

$$\varepsilon(x) = \eta(x) + \lambda \int_a^b R(x,t;\lambda)\,\eta(t)\,\mathrm{d}t,$$

where $R(x,t;\lambda)$ is the unique resolvent kernel corresponding to the given kernel $K(x,t)$.

The estimate

$$|\varepsilon(x)| = |y(x) - \phi(x)| \leq (1 + |\lambda| B_R(\lambda)) B_y$$

follows directly from the integral representation for $\varepsilon(x)$, where

$$B_y = \max_{a \leq x \leq b} |\eta(x)|$$

and $B_R(\lambda)$ is a bound on the integral of the resolvent. Thus, the magnitude of the residual controls the accuracy of the estimation. Since this estimate holds for the approximant $y^*(x)$, the proof is complete. □

Some comments regarding the Method of Collocation may be helpful.

To employ this method, it must be possible to compute the integrals $z_k(x)$, which depend upon the choice of the functions $y_k(x)$. Some typical choices for the functions $y_k(x)$ include various types of polynomials, such as a finite set of powers of x (as illustrated in the example below), Legendre or Chebychev polynomials, or trigonometric polynomials. The best choice depends upon the functional form of the kernel $K(x,t)$.

If the determinant $\det(\mathbf{Y} - \lambda \mathbf{Z}) = 0$, then an alternate choice of nodes and/or basis functions is required. This situation might arise if the basis functions were accidentally chosen to be eigenfunctions of the kernel. If the determinant does not vanish, but is close to zero, then computational difficulties may arise, since the entries in $\mathbf{Y} - \lambda \mathbf{Z}$ might be rather large.

Requiring that $\eta(x_i) = 0$ for all $i = 1, \ldots, n$ is not the only method of producing a small residual. For example, we could require that the coefficients $a_1, \ldots, a_n$ are such that the sum

$$\sum_{i=1}^{n} |\eta(x_i; a_1, \ldots, a_n)|^2$$

is minimal. In this case, the bound on the residual would still control the accuracy of the approximation.

An examination of the graph of the residual $|\eta(x)|$ can lead to improved accuracy for the approximate solution. For if the graph of the residual has a few maxima, but is otherwise small, it is reasonable to suggest that shifting the nodes, i.e., *adapting the mesh*, could lead to a decrease in the magnitude of the residual.

The bound $B_R(\lambda)$ can be chosen to be

$$B_R(\lambda) = \frac{B_K (b-a)}{1 - |\lambda| B_K (b-a)},$$

provided that $|\lambda|$ is less than the radius of convergence of the resolvent series. For other values of $|\lambda|$, an alternate estimate must be made.

A Concise Guide to Computation

- *Preliminary calculations*:
  1. Choose a set $\{x_1,\ldots,x_n\}$ of n nodes with $a \le x_1 < \cdots < x_n \le b$. Usually, the nodes are chosen to be equally spaced, i.e.,
  $$\Delta x_i = x_i - x_{i-1} = \frac{b-a}{n}$$
  for all $i = 1,\ldots,n$.
  2. Choose a set $\{y_1(x),\ldots,y_n(x)\}$ of n functions that are continuous and linearly independent on the interval $[a,b]$.
  3. Compute the elements of the matrices $\mathbf{Y} = (y_{ik})$ and $\mathbf{Z} = (z_{ik})$, where
  $$y_{ik} = y_k(x_i) \quad \text{and} \quad z_{ik} = \int_a^b K(x_i,t)\,y_k(t)\,\mathrm{d}t$$
  for all $i,k = 1,\ldots,n$.
  4. Compute the vector $\mathbf{f} = (f(x_i))$.
- *Determine the approximate solution*:
  1. Compute the determinant $\det(\mathbf{Y} - \lambda\,\mathbf{Z})$. If this determinant vanishes, then adjust the choices of nodes and/or basis functions as necessary.
  2. If the matrix $\mathbf{Y} - \lambda\,\mathbf{Z}$ is invertible, then solve the matrix equation
  $$(\mathbf{Y} - \lambda\,\mathbf{Z})\,\mathbf{a} = \mathbf{f}$$
  for its unique solution
  $$\mathbf{a}^* = (a_1^*,\ldots,a_n^*)^{\mathrm{T}} = (\mathbf{Y} - \lambda\,\mathbf{Z})^{-1}\,\mathbf{f}.$$
  3. Use the solution to formulate the approximation
  $$y^*(x) = \sum_{i=1}^{n} a_i^*\,y_i(x).$$
- *Estimate the accuracy of the approximate solution*:
  1. Formulate the residual $\eta^*(x)$ in terms of $y^*(x)$.
  2. Determine the upper bounds B_{y^*} and $B_K(\lambda)$.
  3. Use these computed bounds to compute an estimate for the upper bound on the possible error $|\varepsilon(x)|$.
  4. If the error must be reduced, then plot the residual to determine an appropriate number of additional nodes and/or basis functions, and repeat the previous steps as appropriate.

Illustrative Examples

- *Example 1*: Consider the Fredholm integral equation

$$\phi(x) = \sin(\pi x) + \lambda \int_0^1 \frac{1}{1+x+t}\,\phi(t)\,\mathrm{d}t.$$

According to the Theorem of Successive Approximation, there are no eigenvalues of the kernel in the disk of radius $1/\|K\|_2 = 1.864419\ldots$, where

$$\|K\|_2 = \left(\int_0^1 \int_0^1 \left(\frac{1}{1+x+t}\right)^2 \mathrm{d}t\,\mathrm{d}x\right)^{1/2} = \sqrt{\ln(\tfrac{4}{3})} = 0.536360\ldots.$$

Therefore, the choice $\lambda = \frac{1}{2}$ is a regular value of the kernel.

Suppose that we desire a polynomial approximation of the sixth degree to the solution $\phi(x)$ of the equation. Following the guide, we choose the seven equally spaced nodes $\{0, \frac{1}{6}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{5}{6}, 1\}$ and the seven basis functions $\{1, x, x^2, x^3, x^4, x^5, x^6\}$. The approximate solution has the form

$$y(x) = \sum_{k=0}^{6} a_k\, y_k(x) = \sum_{k=0}^{6} a_k x^k,$$

and the residual $\eta(x)$ assumes the form

$$\eta(x) = \sum_{k=0}^{6} a_k\,(x^k - \tfrac{1}{2} z_k(x)) - \sin(\pi x).$$

Requiring that $\eta(x_i) = 0$ for $i = 0, \ldots, 6$ leads to the 7×7 matrix system

$$(\mathbf{Y} - \tfrac{1}{2}\,\mathbf{Z})\,\mathbf{a} = \mathbf{f},$$

where $(\mathbf{Y} - \frac{1}{2}\,\mathbf{Z})_{ik} = (x_i^k - \frac{1}{2}\,z_k(x_i))$ and $\mathbf{f} = (0, \frac{1}{2}, \frac{\sqrt{3}}{2}, 1, \frac{\sqrt{3}}{2}, \frac{1}{2}, 0)^{\mathrm{T}}$. With rounding in the last decimal place, the matrix $\mathbf{Y} - \frac{1}{2}\,\mathbf{Z}$ is given by

$$\begin{pmatrix}
+0.653426 & -0.153426 & -0.096574 & -0.070093 & -0.054907 & -0.045093 & -0.038240 \\
+0.690480 & +0.027773 & -0.060180 & -0.059420 & -0.049504 & -0.041217 & -0.035075 \\
+0.720192 & +0.206411 & +0.030342 & -0.021937 & -0.034023 & -0.034061 & -0.031061 \\
+0.744587 & +0.383119 & +0.175321 & +0.070352 & +0.019473 & -0.004209 & -0.014520 \\
+0.764998 & +0.558336 & +0.374995 & +0.245379 & +0.157393 & +0.098583 & +0.059632 \\
+0.782341 & +0.732375 & +0.629535 & +0.531038 & +0.444641 & +0.370834 & +0.308478 \\
+0.797267 & +0.905465 & +0.939070 & +0.955194 & +0.964613 & +0.970775 & +0.975116
\end{pmatrix}.$$

After determining the solution $\mathbf{a}^* = (\mathbf{Y} - \frac{1}{2}\mathbf{Z})^{-1}\mathbf{f}$ to this system, we can formulate the approximate solution

$$\begin{aligned} y^*(x) &= y^*(x; a_1^*, \ldots, a_n^*) \\ &= 0.297340 + 2.932409\,x + 0.124667\,x^2 - 5.092254\,x^3 \\ &\quad - 0.557749\,x^4 + 3.710938\,x^5 - 1.240580\,x^6. \end{aligned}$$

Inspecting the graph of the residual $\eta^*(x)$ shows that $\eta^*(x_i) = 0$ and that $|\eta^*(x)| \leq 0.000036$. With $B_R(\frac{1}{2}) = 2$, the uniform error $|\varepsilon(x)| \leq 0.000072$. Adapting the mesh leads to a slight improvement.

2.6.3 Quadrature Methods

Each of the previously considered numerical methods requires the evaluation of definite integrals in order to produce a function that approximates the solution of a Fredholm integral equation on the interval $[a, b]$. Numerical quadrature methods, on the other hand, require the solution of a system of linear equations in order to do so.

The purpose of any numerical quadrature method is to approximate the definite integral of a continuous function $f(x)$ on a closed interval $[a, b]$ with a finite sum. Regardless of the chosen method, the approximation always takes the form

$$\int_a^b f(x)\,\mathrm{d}x = \sum_{i=1}^{n} w_i\, f(x_i) + E.$$

Each quadrature method requires a predetermined set $\{x_1, \ldots, x_n\}$ of *nodes* with $a \leq x_1 < \ldots < x_n \leq b$ and a set of positive *weights* $\{w_1, \ldots, w_n\}$. The error term E depends upon n, a, b, and the value of some higher derivative of $f(x)$ at an interior point of the interval. The choice of method might depend upon the form of $f(x)$, the interval of integration, the amount of computation required to achieve a given level accuracy, or other factors.

Although there are many such quadrature formulas, they fall into two major classes. Newton–Cotes quadrature formulas require equally spaced nodes, so that

$$\Delta x = x_i - x_{i-1} = \frac{b-a}{n}$$

for all $i = 1, \ldots, n$ and a specific set of weights that do not depend upon the nodes. Formulas of this type include the Trapezoid rule, Simpson's rule, and Bode's rule. On the other hand, the nodes in Gaussian quadrature formulas are chosen to be the zeroes of some orthogonal polynomial of degree n, and the weights are given in terms of these polynomials and/or their derivatives evaluated at the nodes. Common

choices are Legendre, Chebychev, or Jacobi polynomials. (Laguerre polynomials are used if the interval is $[0,+\infty)$, and Hermite polynomials are used if the interval is $(-\infty,+\infty)$.)

Now consider the integral equation

$$\phi(x) = f(x) + \lambda \int_a^b K(x,t)\,\phi(t)\,\mathrm{d}t,$$

where $f(x)$ is continuous on the interval $[a,b]$ and the kernel $K(x,t)$ is continuous on the square $Q(a,b)$. After choosing an appropriate numerical method, we substitute each node x_i into the integral equation to obtain the n equations

$$\phi(x_i) = f(x_i) + \lambda \int_a^b K(x_i,t)\,\phi(t)\,\mathrm{d}t.$$

By substituting the nodes and positive weights indicated for the chosen method, the value of each of these integrals can be expressed in the form

$$\int_a^b K(x_i,t)\,\phi(t)\,\mathrm{d}t = \sum_{j=1}^{n} w_j\,K(x_i,x_j)\,\phi(x_j) + E(x_i).$$

Replacing the definite integrals with the finite sums produces the n equations

$$\phi(x_i) = f(x_i) + \lambda \sum_{j=1}^{n} w_j\,K(x_i,x_j)\,\phi(x_j) + \lambda\,E(x_i).$$

After discarding the error term, we arrive at the system

$$y_i = f(x_i) + \lambda \sum_{j=1}^{n} w_j\,K(x_i,x_j)\,y_j$$

of n equations in the n unknowns y_i. It was necessary to replace the *exact* values $\phi(x_i)$ with the *approximate* values y_i since the error terms were discarded. If the $E(x_i)$ is small, then y_i is close to $\phi(x_i)$. In matrix form, this linear system becomes

$$(\mathbf{I} - \lambda\,\mathbf{KW})\,\mathbf{y} = \mathbf{f}.$$

In this matrix equation, we have set $\mathbf{K} = (K_{ij}) = (K(x_i,x_j))$, $\mathbf{y} = (y_1,\ldots,y_n)^{\mathrm{T}}$, and $\mathbf{f} = (f(x_1),\ldots,f(x_n))^{\mathrm{T}}$. The matrix $\mathbf{W} = (W_{ij})$ is a diagonal matrix with weights w_i appearing on the diagonal, i.e., $W_{ii} = w_i$ and $W_{ij} = 0$, if $i \neq j$. Assuming that the matrix $(\mathbf{I} - \lambda\,\mathbf{KW})$ is invertible, the solution assumes the form

$$\mathbf{y} = (\mathbf{I} - \lambda\,\mathbf{KW})^{-1}\,\mathbf{f}.$$

Once $\mathbf{y} = (y_1,\ldots,y_n)^{\mathrm{T}}$ is determined, a continuous interpolating function $y(x)$ can be constructed on the interval $[a,b]$ that passes through all of the points (x_i,y_i).

A simple way to accomplish this construction would be to define a continuous interpolating function by setting

$$y(x) = f(x) + \lambda \sum_{j=1}^{n} w_j K(x,x_j) y_j.$$

Another simple way to produce a continuous interpolation function would be to use the classical Lagrange interpolation formula

$$y(x) = \sum_{i=1}^{n} y_i \left(\prod_{\substack{1 \le k \le n \\ k \ne i}} \frac{x - x_k}{x_i - x_k} \right)$$

directly when n is small. For example, if $n = 3$, then

$$\begin{aligned} y(x) = {} & y_1 \left(\frac{x-x_2}{x_1-x_2} \right) \left(\frac{x-x_3}{x_1-x_3} \right) \\ & + y_2 \left(\frac{x-x_1}{x_2-x_1} \right) \left(\frac{x-x_3}{x_2-x_3} \right) \\ & + y_3 \left(\frac{x-x_1}{x_3-x_1} \right) \left(\frac{x-x_2}{x_3-x_2} \right). \end{aligned}$$

By inspection, it is clear that $y(x_1) = y_1$, $y(x_2) = y_2$, and $y(x_3) = y_3$.

However, if n is large, then essentially unpredictable anomalies may arise in the interpolation function. For example, if $n = 50$, then it is possible to construct a polynomial $p(x)$ of degree 49 with the Lagrange formula so that $y_i = p(x_i)$ for all $i = 1, \ldots, 50$. Its derivative $p'(x)$ might have as many as 48 simple zeroes, at which as many as 48 relative extrema may occur. If a large number of these zeroes fall between two successive nodes, then it is likely that "polynomial wiggles" would be introduced into the graph of the interpolating polynomial between those two nodes. As a consequence, $p(x)$ would then approximate the solution poorly between successive nodes.

To avoid the introduction of polynomial wiggles, *natural cubic splines* are often used in practice. In a sense, cubic splines are the smoothest possible interpolating functions.

Briefly, a cubic spline is a piecewise cubic polynomial. After a choice of nodes $\{x_0, x_1, \ldots, x_n\}$ is made, with $a = x_0 < x_1 \ldots < x_n = b$, a cubic polynomial $q_i(x)$ is defined on each subinterval $[x_{i-1}, x_i]$, for $i = 1, \ldots, n$. These cubics are defined uniquely by imposing continuity conditions on each $q_i(x)$ and its first two derivatives at the interior nodes and requiring that $q''(a) = q''(b) = 0$ at the endpoints. To be more specific, for each $i = 1, \ldots, n-1$, we require that $q_i(x_i) = q_{i+1}(x_i)$, $q_i'(x_i) = q_{i+1}'(x_i)$ and that $q_i''(x_i) = q_{i+1}''(x_i)$. Since the values of the adjacent cubics and their first two derivatives match at the nodes, the interpolation is quite smooth. Furthermore, since each $q_i(x)$ can have no more than two critical points in each subinterval, polynomial wiggles cannot arise.

Illustrative Examples

- *Example 1*: Consider the inhomogeneous integral equation

$$\phi(x) = x^3 + \frac{1}{2}\int_{-1}^{+1} \sin\left(\frac{\pi xt}{2}\right)\phi(t)\,\mathrm{d}t. \tag{2.30}$$

After replacing the kernel with a separable one derived from fifth partial sum of the Maclaurin series for $\sin x$, we can determine an approximation $y(x)$ to $\phi(x)$ by solving the equation

$$y(x) = x^3 + \frac{1}{2}\int_{-1}^{+1}\left(\frac{1}{2}\pi xt - \frac{1}{48}\pi^3x^3t^3 + \frac{1}{3840}\pi^5x^5t^5\right)y(t)\,\mathrm{d}t.$$

By employing the method of Sect. 1.3, we obtain

$$y(x) = 0.565620x + 0.847692x^3 + 0.014047x^5.$$

This approximation is quite accurate in the sense that the residual

$$\begin{aligned}\eta_y(x) &= y(x) - x^3 - \frac{1}{2}\int_{-1}^{+1}\sin\left(\frac{\pi xt}{2}\right)y(t)\,\mathrm{d}t\\ &\approx 0.000660x^7 - 0.000018x^9.\end{aligned}$$

(a) An application of Simpson's rule yields another approximation $z(x)$ to $\phi(x)$. If we partition the interval $[-1,+1]$ into six subintervals, then the length of each subinterval is $\Delta x = 1/3$. The equally distributed nodes are

$$\{x_1,\ldots,x_7\} = \{-1,-\tfrac{2}{3},-\tfrac{1}{3},0,+\tfrac{1}{3},+\tfrac{2}{3},+1\},$$

and the corresponding weights are

$$\{w_1,\ldots,w_7\} = \{\tfrac{1}{9},\tfrac{4}{9},\tfrac{2}{9},\tfrac{4}{9},\tfrac{2}{9},\tfrac{4}{9},\tfrac{1}{9}\}.$$

If we set $x = x_i$ in Eq. (2.30), then we obtain

$$\phi(x_i) = x_i^3 + \frac{1}{2}\int_{-1}^{+1}\sin\left(\frac{\pi x_i t}{2}\right)\phi(t)\,\mathrm{d}t$$

for each $i = 1,\ldots,7$. Approximating each integral with Simpson's Rule yields the approximate equations

$$z(x_i) = x_i^3 + \frac{1}{2}\sum_{j=1}^{7} w_j \sin\left(\frac{\pi x_i x_j}{2}\right)z(x_j).$$

Now set $\mathbf{f} = (-1, -\frac{8}{27}, -\frac{1}{27}, 0, +\frac{1}{27}, +\frac{8}{27}, +1)^{\mathrm{T}}$, and solve the linear system

$$(\mathbf{I} - \tfrac{1}{2}\mathbf{KW})\mathbf{z} = \mathbf{f},$$

where $\mathbf{K} = (K_{ij}) = (\sin(\pi x_i x_j/2))$ and $\mathbf{W} = (W_{ij}) = w_i\,\delta_{ij}$, to obtain

$$\mathbf{z} = \begin{pmatrix} -1.425492 \\ -0.630291 \\ -0.220551 \\ 0.000000 \\ +0.220551 \\ +0.630291 \\ +1.425492 \end{pmatrix}.$$

After substituting these values into the interpolation formula, we obtain the approximate solution

$$\begin{aligned} z(x) &= x^3 + \frac{1}{2}\sum_{j=1}^{7} w_j \sin\left(\frac{\pi x x_j}{2}\right) z(x_j) \\ &\approx 0.567809\,x + 0.842899\,x^3 + 0.015578\,x^5. \end{aligned}$$

In this case, the residual is

$$\begin{aligned} \eta_z(x) &= z(x) - x^3 - \frac{1}{2}\int_{-1}^{+1} \sin\left(\frac{\pi x t}{2}\right) z(t)\,\mathrm{d}t \\ &\approx 0.0023436\,x - 0.004890\,x^3 + 0.001542\,x^5 \\ &\quad - 0.000158\,x^7 + 0.000007\,x^9. \end{aligned}$$

(b) A comparable approximation to the solution $\phi(x)$ can be had by employing Gaussian quadrature.

The Legendre polynomial of degree seven is

$$P_7(x) = \frac{1}{16}\left(429x^7 - 693x^5 + 315x^3 - 35x\right).$$

With this method, the nodes $\{u_1, \ldots, u_7\}$ are the seven roots of $P_7(x)$

$$\{-0.949107, -0.741531, -0.405845, 0, +0.405845, +0.741531, +0.949107\},$$

and the corresponding weights $\{v_1, \ldots, v_7\}$ are computed from the formula

$$v_i = \frac{2}{(1-u_i^2)\,(P_7'(u_i))^2}$$

to be

$$\{0.129484, 0.279705, 0.381830, 0.417959, 0.381830, 0.279705, 0.129484\}.$$

If we set $x = u_i$ in Eq. (2.30), then we obtain

$$\phi(u_i) = u_i^3 + \frac{1}{2}\int_{-1}^{+1} \sin\left(\frac{\pi u_i t}{2}\right)\phi(t)\,\mathrm{d}t$$

for each $i = 1, \ldots, 7$. Approximating each of these integrals by Gaussian quadrature yields the approximate equations

$$g(u_i) = u_i^3 + \frac{1}{2}\sum_{j=1}^{7} v_j \sin\left(\frac{\pi u_i u_j}{2}\right) g(u_j).$$

Now set $\mathbf{h} = (u_1^3, \ldots, u_7^3)^{\mathrm{T}}$ and solve the linear system

$$(\mathbf{I} - \tfrac{1}{2}\mathbf{KV})\mathbf{s} = \mathbf{h},$$

where $\mathbf{K} = (K_{ij}) = \sin(\pi u_i u_j/2)$ and $\mathbf{V} = (V_{ij}) = v_i\,\delta_{ij}$, to obtain

$$\mathbf{s} = \begin{pmatrix} -1.271807 \\ -0.768011 \\ -0.286296 \\ 0.000000 \\ +0.286296 \\ +0.768011 \\ +1.271807 \end{pmatrix}.$$

After substituting these values into the interpolation formula, we obtain

$$\begin{aligned} g(x) &= x^3 + \frac{1}{2}\sum_{j=1}^{7} v_j \sin\left(\frac{\pi x u_j}{2}\right) g(u_j) \\ &= 0.565420\,x + 0.847750\,x^3 + 0.014041\,x^5. \end{aligned}$$

The approximation $g(x)$ is very close to $\phi(x)$, since the residual

$$\eta_g(x) = g(x) - x^3 - \frac{1}{2}\int_{-1}^{+1} \sin\left(\frac{\pi x t}{2}\right) g(t)$$
$$\approx 10^{-8} \times (-4.337x + 1.543x^3 - 0.181x^5).$$

The approximation $g(x)$ is also very close to $y(x)$ determined above as the solution to the integral equation with the separable kernel.

The graphs of the approximations $y(x)$, $z(x)$, and $g(x)$ are "within the ink" of each other, i.e., if all three approximations are plotted in the same plane, they are indistinguishable.

The number of nodes in the illustrative examples was deliberately chosen to be small in order to avoid unnecessary numerical complications. It should be noted that significant problems relating to stability may arise if the number of nodes and/or the number of basis functions is increased.

Chapter 3
Fredholm Integral Equations of the Second Kind (Hermitian Kernel)

A *Hermitian kernel* is a kernel that satisfies the property

$$K^*(x,t) = \overline{K(t,x)} = K(x,t)$$

in the square $Q(a,b) = \{(x,t)\colon a \le x \le b \text{ and } a \le t \le b\}$. We assume as usual that $K(x,t)$ is continuous in $Q(a,b)$.

In this chapter, we explore the significant and astoundingly beautiful role played by Hermitian kernels within the theory of Fredholm integral equations of the second kind.

In Sect. 3.1, we present several tools of the trade that are indispensible for the comprehension of the material in the chapter.

In Sect. 3.2, we explore the properties of Hermitian kernels, as well as their eigenvalues and eigenfunctions.

In Sect. 3.3, we unveil the structure of Hermitian kernels with a finite number of eigenvalues, and then state expansion theorems for them and their iterations. We construct the resulting resolvent kernel in order to be able to solve integral equations with Hermitian separable kernels.

In Sect. 3.4, we will extend these results to Hermitian kernels with an infinite number of eigenvalues via Fourier series. Here, the Hilbert–Schmidt Expansion Theorem is the key result in this endeavor.

In Sect. 3.5, we present some numerical methods for computing eigenvalues and eigenfunctions of Hermitian kernels.

3.1 Tools of the Trade

In this chapter, the reader should be familiar with the following topics:

- *Riemann integrable functions*: In this chapter, the term "integrable" means *Riemann integrable*. The space consisting of all complex-valued Riemann

S.M. Zemyan, *The Classical Theory of Integral Equations: A Concise Treatment*,
DOI 10.1007/978-0-8176-8349-8_3,

integrable functions on the interval $[a,b]$ will be denoted by $\mathcal{R}[a,b]$. The inner product space that consists of all square-integrable functions $f(x)$ for which $\|f\|_2 < +\infty$ on $[a,b]$ will be denoted by $\mathcal{R}^2[a,b]$.

- *Mutually orthonormal systems of functions*: A set $\{\phi_n(x)\}$ of nonzero complex-valued continuous functions defined on the interval $[a,b]$ is called a *mutually orthogonal system of functions* if

$$\langle \phi_n, \phi_m \rangle = \int_a^b \phi_n(x)\,\overline{\phi_m(x)}\,\mathrm{d}x = 0$$

whenever $n \neq m$. If, in addition, it is true that

$$\|\phi_n\|_2^2 = \langle \phi_n, \phi_n \rangle = \int_a^b |\phi_n(x)|^2\,\mathrm{d}x = 1$$

for every n, then the set $\{\phi_n(x)\}$ is called a *mutually orthonormal system of functions*. These systems may be either finite or infinite.

For example, if $\psi_n(x) = \mathrm{e}^{inx}$, then the set $\{\psi_n(x)\}_{n=-\infty}^{\infty}$ is a mutually orthogonal system on the interval $[0,2\pi]$. Since $\|\psi_n(x)\|_2 = \sqrt{2\pi}$, the set $\{\phi_n(x)\}$ is a mutually orthonormal system, where $\phi_n(x) = \psi_n(x)/\|\psi_n\|_2$.

The set of Legendre polynomials $\{P_n(x)\}_{n=0}^{\infty}$ also constitutes a mutually orthogonal system on the interval $[-1,+1]$, for if $n \neq m$, then

$$\int_{-1}^{+1} P_n(x)\,P_m(x)\,\mathrm{d}x = 0.$$

Since

$$\|P_n\|_2^2 = \int_{-1}^{+1} P_n^2(x)\,\mathrm{d}x = \frac{2}{2n+1}$$

for every $n \geq 0$, the set $\{L_n(x)\}_{n=0}^{\infty}$ of functions, where

$$L_n(x) = \sqrt{\frac{2n+1}{2}}\,P_n(x),$$

constitutes a mutually orthonormal system on the interval $[-1,+1]$.

Every mutually orthonormal system of functions $\{\phi_n(x)\}$ must necessarily be a linearly independent set on its interval of definition. For if

$$c_1\,\phi_1(x) + \cdots + c_k\,\phi_k(x) = 0,$$

then $c_i\,(\phi_i,\phi_i) = c_i = 0$ for each value of $i = 1,\ldots,k$.

A function $f(x)$ belongs to the *span* of the mutually orthonormal system $\{\phi_1,\ldots,\phi_k\}$ if and only if there exist constants c_i such that

$$f(x) = c_1\,\phi_1(x) + \cdots + c_k\,\phi_k(x).$$

However, since the property of orthonormality implies that $c_i = \langle f, \phi_i \rangle$, any such $f(x)$ can always be represented as a linear combination

$$f(x) = \langle f, \phi_1 \rangle \, \phi_1(x) + \cdots + \langle f, \phi_k \rangle \, \phi_k(x).$$

In this chapter, it will be important to know that any linearly independent set of continuous functions can be orthonormalized via the *Gram–Schmidt Orthonormalization Procedure*. Let $\{\psi_n(x)\}$ be a linearly independent set of nonzero, complex-valued continuous functions defined on the interval $[a,b]$. If we set

$$\phi_1(x) = \frac{\psi_1(x)}{\|\psi_1\|_2},$$

then $\|\phi_1\|_2 = 1$. If we set

$$v_2(x) = \psi_2(x) - \langle \psi_2, \phi_1 \rangle \, \phi_1(x),$$

then

$$\langle v_2, \phi_1 \rangle = \langle \psi_2, \phi_1 \rangle - \langle \psi_2, \phi_1 \rangle \, \langle \phi_1, \phi_1 \rangle = 0,$$

so that v_2 is orthogonal to ϕ_1. If we set

$$\phi_2(x) = \frac{v_2(x)}{\|v_2\|_2},$$

then ϕ_2 is orthogonal to ϕ_1 and $\|\phi_2\| = 1$.

Continuing in this manner, if we set

$$v_k(x) = \psi_k(x) - \sum_{i=1}^{k-1} \langle \psi_k, \phi_i \rangle \, \phi_i(x),$$

then

$$\langle v_k, \phi_j \rangle = \langle \psi_k, \phi_j \rangle - \sum_{i=1}^{k-1} \langle \psi_k, \phi_i \rangle \, \langle \phi_i, \phi_j \rangle = 0,$$

so that $v_k(x)$ is orthogonal to ϕ_j for all $j = 1, \ldots, k-1$. If we set

$$\phi_k(x) = \frac{v_k(x)}{\|v_k\|_2},$$

then ϕ_k is orthogonal to ϕ_j for all $j = 1, \ldots, k-1$ and $\|\phi_k\| = 1$. The set $\{\phi_1, \ldots, \phi_k\}$ is a mutually orthonormal system.

To illustrate this procedure, we apply it to the set $\{1, x, x^2, x^3\}$, which is linearly independent on the interval $[4,8]$, to construct the mutually orthonormal system $\{\phi_1(x), \phi_2(x), \phi_3(x), \phi_4(x)\}$ on $[4,8]$. Since

$$\|\psi_1\|_2^2 = \langle 1, 1 \rangle = \int_4^8 1 \cdot 1 \, \mathrm{d}x = 4,$$

we set $\phi_1(x) = \frac{1}{2}$. Next, we set

$$v_2(x) = x - \langle x, \phi_1 \rangle \, \phi_1 = x - 6.$$

Since

$$\|v_2\|_2^2 = \int_4^8 (x-6)^2 \, \mathrm{d}x = \frac{16}{3},$$

we set

$$\phi_2(x) = \frac{\sqrt{3}}{4}(x-6).$$

Next, we set

$$v_3(x) = x^2 - \langle x^2, \phi_1 \rangle \, \phi_1(x) - \langle x^2, \phi_2 \rangle \, \phi_2(x) = x^2 - 12x + \frac{104}{3}.$$

Since

$$\|v_3\|_2^2 = \int_4^8 \left(x^2 - 12x + \frac{104}{3} \right)^2 \mathrm{d}x = \frac{256}{45},$$

we set

$$\phi_3(x) = \frac{\sqrt{5}}{16}(3x^2 - 36x + 104).$$

Next, we set

$$\begin{aligned} v_4(x) &= x^3 - \langle x^3, \phi_1 \rangle \, \phi_1(x) - \langle x^3, \phi_2 \rangle \, \phi_2(x) - \langle x^3, \phi_3 \rangle \, \phi_3(x) \\ &= x^3 - 18x^2 + \frac{528}{5}x - \frac{1008}{5}. \end{aligned}$$

Since

$$\|v_4\|_2^2 = \int_4^8 \left(x^3 - 18x^2 + \frac{528}{5}x - \frac{1008}{5} \right)^2 \mathrm{d}x = \frac{1024}{175},$$

we set

$$\phi_4(x) = \frac{\sqrt{7}}{32}\left(5x^3 - 90x^2 + 528x - 1008\right).$$

Since $\langle \phi_n, \phi_m \rangle = 0$ if $n \neq m$ and $\langle \phi_n, \phi_n \rangle = 1$ for each n, the resulting set $\{\phi_1, \phi_2, \phi_3, \phi_4\}$ constitutes a mutually orthonormal system on $[4, 8]$.

- *Fredholm operators*: The operator $\mathbf{K}$ defined by

$$\mathbf{K}\colon \phi \in \mathcal{R}^2[a,b] \to \mathbf{K}\phi = \int_a^b K(x,t)\,\phi(t)\,\mathrm{d}t \in \mathcal{R}^2[a,b]$$

is called the *Fredholm operator*. Fredholm operators are always linear, since $\mathbf{K}(\phi_1 + \phi_2) = \mathbf{K}\phi_1 + \mathbf{K}\phi_2$ and $\mathbf{K}(\alpha\phi) = \alpha\mathbf{K}\phi$ for any complex constant α.

Two Fredholm operators, $\mathbf{K}$ and $\mathbf{L}$, are said to be *equal* if $\mathbf{K}\phi = \mathbf{L}\phi$ for all $\phi \in \mathcal{R}^2[a,b]$.

Fredholm operators can be composed to define new Fredholm operators. Suppose that $K(x,t)$ and $L(x,t)$ are complex-valued kernels defined on the square $Q(a,b)$ and integrable with respect to each of their variables. If we define the composition $M(x,t)$ of these two kernels to be

$$M(x,t) = \int_a^b K(x,s)\,L(s,t)\,\mathrm{d}s,$$

then for every $\phi \in \mathcal{R}^2[a,b]$, we have

$$\begin{aligned}
\mathbf{M}\phi &= \int_a^b M(x,s)\,\phi(s)\,\mathrm{d}s \\
&= \int_a^b \left(\int_a^b K(x,t)\,L(t,s)\,\mathrm{d}t \right) \phi(s)\,\mathrm{d}s \\
&= \int_a^b K(x,t) \left(\int_a^b L(t,s)\,\phi(s)\,\mathrm{d}s \right) \mathrm{d}t \\
&= \mathbf{KL}\phi.
\end{aligned}$$

If $\mathbf{K} = \mathbf{L}$, then we write $\mathbf{K}^2\phi = \mathbf{K}(\mathbf{K}\phi)$ and $\mathbf{K}^n\phi = \mathbf{K}(\mathbf{K}^{n-1}\phi)$. Thus, iterated operators can naturally be expressed in terms of iterated kernels. Specifically,

$$\mathbf{K}^m\phi = \int_a^b K_m(x,t)\,\phi(t)\,\mathrm{d}t.$$

The composition of two Fredholm operators $\mathbf{K}$ and $\mathbf{L}$ is generally not commutative, but it is associative.

A set $\mathcal{F} \subseteqq \mathcal{R}^2[a,b]$ is called *bounded* if there exists a constant B such that $\|f\|_2 < B$ for all $f \in \mathcal{F}$. A set $\mathcal{F} \subseteqq \mathcal{R}^2[a,b]$ is called *compact* if a convergent subsequence can be selected from any infinite sequence in $\mathcal{F}$.

Proposition 3.1.1. *If $K(x,t)$ is a complex-valued kernel whose norm $\|K\|_2$ is finite, then the corresponding Fredholm operator $\mathbf{K}$ maps any bounded set in $\mathcal{R}^2[a,b]$ into a compact set in $\mathcal{R}^2[a,b]$.*

A linear operator $\mathbf{K}$ is *bounded* if there exists a constant $C \geq 0$ such that $\|\mathbf{K}\phi\|_2 \leq C\,\|\phi\|_2$ for all $\phi \in \mathcal{R}^2[a,b]$. The *norm* $\|\mathbf{K}\|$ *of the bounded linear operator* $\mathbf{K}$ is the smallest possible value of C for which this inequality holds. Precisely,

$$\|\mathbf{K}\| = \sup \left\{ \frac{\|\mathbf{K}\phi\|_2}{\|\phi\|_2} \colon \phi \in \mathcal{R}^2[a,b],\ \phi \neq 0 \right\}.$$

If $\mathbf{K}$ is bounded, then the inequality $\|\mathbf{K}\phi\|_2 \leq \|\mathbf{K}\| \cdot \|\phi\|_2$ holds for all $\phi \in \mathcal{R}^2[a,b]$.

Bounded operators satisfy the triangle inequality. For if $\phi \in \mathcal{R}^2[a,b]$ and $\mathbf{K}$ and $\mathbf{L}$ are any two bounded operators, then

$$\begin{aligned}\|(\mathbf{K}+\mathbf{L})\phi\|_2 &= \|\mathbf{K}\phi+\mathbf{L}\phi\|_2\\ &\le \|\mathbf{K}\phi\|_2+\|\mathbf{L}\phi\|_2\\ &\le (\|\mathbf{K}\|+\|\mathbf{L}\|)\cdot\|\phi\|_2,\end{aligned}$$

from which we immediately conclude that

$$\|\mathbf{K}+\mathbf{L}\| \le \|\mathbf{K}\|+\|\mathbf{L}\|.$$

If $K(x,t)$ is a kernel for which $\|K\|_2 < +\infty$ and $\mathbf{K}$ is the corresponding Fredholm operator, then $\mathbf{K}$ is bounded and $\|\mathbf{K}\| \le \|K\|_2$. An application of the Cauchy–Schwarz inequality gives

$$|\mathbf{K}\phi(x)|^2 \le \left(\int_a^b |K(x,s)|^2\,\mathrm{d}s\right)\left(\int_a^b |\phi(s)|^2\,\mathrm{d}s\right).$$

Upon integration, we have

$$\begin{aligned}\|\mathbf{K}\phi\|_2^2 &= \int_a^b |\mathbf{K}\phi(x)|^2\,\mathrm{d}x\\ &\le \left(\int_a^b\int_a^b |K(x,s)|^2\,\mathrm{d}s\,\mathrm{d}x\right)\left(\int_a^b |\phi(s)|^2\,\mathrm{d}s\right)\\ &= \|K\|_2^2\,\|\phi\|_2^2\end{aligned}$$

from which the conclusion follows directly.

If $\mathbf{K}$ and $\mathbf{L}$ are two bounded operators, then the composition $\mathbf{KL}$ is also a bounded operator. Since

$$\begin{aligned}\|\mathbf{KL}\phi\|_2 &\le \|\mathbf{K}\|\cdot\|\mathbf{L}\phi\|_2\\ &\le \|\mathbf{K}\|\cdot\|\mathbf{L}\|\cdot\|\phi\|_2,\end{aligned}$$

it follows immediately that $\|\mathbf{KL}\| \le \|\mathbf{L}\|\cdot\|\mathbf{K}\|$.

If we set $K^*(x,t) = \overline{K(t,x)}$, then the operator defined by

$$\mathbf{K}^*\colon\ \phi \in \mathcal{R}^2[a,b] \to \mathbf{K}^*\phi(x) = \int_a^b K^*(x,t)\,\phi(t)\,\mathrm{d}t \in \mathcal{R}^2[a,b]$$

is called the *conjugate transpose Fredholm operator*. Clearly, $\mathbf{K}^{**} = \mathbf{K}$, since $K^{**}(x,t) = K(x,t)$.

Fredholm operators and their conjugates satisfy the relation

$$\langle \mathbf{K}\phi, \psi \rangle = \langle \phi, \mathbf{K}^* \psi \rangle$$

for all $\phi, \psi \in \mathcal{R}^2[a,b]$. To see this, we interchange the order and variables of integration to obtain

$$\begin{aligned}
\langle \mathbf{K}\phi, \psi \rangle &= \int_a^b \left(\int_a^b K(x,t)\, \phi(t)\, \mathrm{d}t \right) \overline{\psi(x)}\, \mathrm{d}x \\
&= \int_a^b \left(\int_a^b K(x,t)\, \overline{\psi(x)}\, \mathrm{d}x \right) \phi(t)\, \mathrm{d}t \\
&= \int_a^b \left(\int_a^b K(t,x)\, \overline{\psi(t)}\, \mathrm{d}t \right) \phi(x)\, \mathrm{d}x \\
&= \int_a^b \phi(x) \overline{\left(\int_a^b \overline{K(t,x)}\, \psi(t)\, \mathrm{d}t \right)}\, \mathrm{d}x \\
&= \langle \phi, \mathbf{K}^* \psi \rangle.
\end{aligned}$$

The conjugate of a composition of two Fredholm operators is equal to the composition of their conjugates in reverse order, i.e.,

$$(\mathbf{KL})^* \phi = \mathbf{L}^* \mathbf{K}^* \phi$$

for all $\phi \in \mathcal{R}^2[a,b]$. If $\phi, \psi \in \mathcal{R}^2[a,b]$, then

$$\langle (\mathbf{KL})\psi, \phi \rangle = \langle \psi, (\mathbf{KL})^* \phi \rangle.$$

On the other hand,

$$\langle \mathbf{KL}\psi, \phi \rangle = \langle \mathbf{L}\psi, \mathbf{K}^* \phi \rangle = \langle \psi, \mathbf{L}^* \mathbf{K}^* \phi \rangle.$$

By subtracting these last two equations, we obtain

$$\langle \psi, (\mathbf{KL})^* \phi - \mathbf{L}^* \mathbf{K}^* \phi \rangle = 0.$$

If we choose ψ to be the second factor of this inner product, then we obtain

$$\| (\mathbf{KL})^* \phi - \mathbf{L}^* \mathbf{K}^* \phi \|^2 = 0$$

for all $\phi \in \mathcal{R}^2[a,b]$, from which the conclusion follows immediately. In particular, if we choose $\mathbf{K} = \mathbf{L}$, then we obtain $(\mathbf{K}^2)^* = (\mathbf{K}^*)^2$. A short inductive argument then shows that $(\mathbf{K}^m)^* = (\mathbf{K}^*)^m$ for all $m \geq 1$.

- *Convergence in the mean*: If $f,g \in \mathcal{R}^2[a,b]$, then the *mean square distance* between them is defined as

$$\|f-g\|_2 = \left(\int_a^b |f(x)-g(x)|^2 \, \mathrm{d}x \right)^{1/2}.$$

A sequence $\{f_n(x)\}$ in $\mathcal{R}^2[a,b]$ *converges in the mean to a limit function* $f \in \mathcal{R}^2[a,b]$ if the mean square distance $\|f-f_n\|_2 \to 0$ as $n \to +\infty$.

An important criterion (due to Riesz and Fischer) for convergence in the mean is the following: A sequence $\{f_n(x)\}$ in $\mathcal{R}^2[a,b]$ converges in the mean to a function $f(x) \in \mathcal{R}^2[a,b]$ if and only if

$$\lim_{m,n\to\infty} \|f_n - f_m\|_2 \to 0.$$

Let $\{f_n(x)\}$ be a sequence of functions in $\mathcal{R}^2[a,b]$ and define its *sequence of partial sums* as

$$s_N(x) = \sum_{n=1}^{N} f_n(x).$$

If the sequence $\{s_N(x)\}$ of partial sums converges in the mean to a function $f(x) \in \mathcal{R}^2[a,b]$, i.e., if $\|f-s_N\|_2 \to 0$ as $N \to \infty$, then we say that $f(x)$ is the *sum of the infinite series* and we write

$$f(x) = \sum_{n=1}^{\infty} f_n(x).$$

If a series converges in the mean, then it can be integrated term-by-term. More precisely, if $g \in \mathcal{R}^2[a,b]$, then

$$\int_a^b f(x)\, g(x)\, \mathrm{d}x = \sum_{n=1}^{\infty} \int_a^b f_n(x)\, g(x)\, \mathrm{d}x.$$

It is easy to show that this is true. Since

$$|\langle f-s_N, \overline{g} \rangle| \le \|f-s_N\|_2 \cdot \|g\|_2 \to 0,$$

it follows that

$$\begin{aligned} \langle f, \overline{g} \rangle &= \lim_{N\to\infty} \langle s_N, \overline{g} \rangle \\ &= \lim_{N\to\infty} \sum_{n=1}^{N} \langle f_n, \overline{g} \rangle \\ &= \sum_{n=1}^{\infty} \langle f_n, \overline{g} \rangle. \end{aligned}$$

- *Completeness and Fourier series*: Let $\Phi = \{\phi_n(x)\}_{n=1}^{\infty}$ be a mutually orthonormal system of functions, and let Φ_m be a finite subset of Φ.

  We have previously noted that if $f \in \text{span}\{\Phi_m\}$, then $f(x)$ will have the representation

$$f(x) = \langle f, \phi_1 \rangle \, \phi_1(x) + \cdots + \langle f, \phi_m \rangle \, \phi_m(x).$$

  If $f \notin \text{span}\{\Phi_m\}$, then the mean square distance $d(f, \sigma_m)$ between $f(x)$ and any element $\sigma_m(x) \in \text{span}\{\Phi_m\}$ is given by

$$\begin{aligned} d^2(f, \sigma_m) &= \left\| f - \sum_{n=1}^{m} c_n \phi_n \right\|_2^2 \\ &= \|f\|_2^2 + \sum_{n=1}^{m} |c_n - \langle f, \phi_n \rangle|^2 - \sum_{n=1}^{m} |\langle f, \phi_n \rangle|^2, \end{aligned}$$

  having assumed a priori that $f \in \mathcal{R}^2[a,b]$. It is clear that $d(f, \sigma_m)$ is minimized by choosing $c_n = \langle f, \phi_n \rangle$ for every $n = 1, \ldots, m$. The constants $\langle f, \phi_n \rangle$ are called the *Fourier coefficients of* $f(x)$. For these choices, we have

$$d^2(f, \Phi_m) = \min_{\Phi_m} d^2(f, \sigma_m) = \|f\|_2^2 - \sum_{n=1}^{m} |\langle f, \phi_n \rangle|^2,$$

  from which we immediately deduce that the coefficient inequality

$$\sum_{n=1}^{m} |\langle f, \phi_n \rangle|^2 \leq \|f\|_2^2$$

  holds for every $m \geq 1$. As $m \to \infty$, the minimal distances $d(f, \Phi_m)$ decrease to the nonnegative limit

$$d(f, \Phi) = \lim_{m \to \infty} d(f, \Phi_m) = \|f\|_2^2 - \sum_{n=1}^{\infty} |\langle f, \phi_n \rangle|^2,$$

  and consequently, the coefficient inequality becomes

$$\sum_{n=1}^{\infty} |\langle f, \phi_n \rangle|^2 \leq \|f\|_2^2.$$

  This classical result is known as *Bessel's inequality*. Equality holds here if and only if $d(f, \Phi) = 0$; the resulting equality is known as *Parseval's identity* or *Parseval's equation*.

  Suppose that $f \in \mathcal{R}^2[a,b]$ and that $\langle f, \phi \rangle = 0$ for all $\phi \in \Phi$. If it must be the case that $\|f\|_2 = 0$ for every such f, then the system Φ is said to be *complete*; if there exists such an f for which $\|f\|_2 > 0$, then the system Φ is said to be *incomplete*.

Most of the mutually orthogonal systems of functions to be encountered will consist of continuous functions. For example, the system

$$\Phi = \{\sin nx, \cos nx\}_{n=1}^{\infty}$$

is incomplete in $\mathcal{R}^2[0,2\pi]$ because $\langle 1, \sin nx\rangle = \langle 1, \cos nx\rangle = 0$ for every $n \geq 1$, but $\|1\|_2 = \sqrt{2\pi}$. However, it can be shown that the augmented orthogonal system $\Phi \cup \{1\}$ is complete.

Suppose that Φ is a mutually orthonormal system of continuous functions on the interval $[a,b]$. Given $f \in \mathcal{R}^2[a,b]$, we can always compute the values $\langle f, \phi_n\rangle$ and then use them to construct the series

$$f(x) \sim \sum_{n=1}^{\infty} \langle f, \phi_n\rangle \, \phi_n(x).$$

We use the symbol $\sim$ to indicate only that the corresponding series was generated from $f(x)$. This series is called the *Fourier series* of $f(x)$.

As with any series, two basic questions arise here:

– Does the Fourier series converge?
– If it does converge, does it converge to $f(x)$ as we might expect?

To address the first question, we note that varied criteria exist for the convergence of a Fourier series. For example, if $f(x)$ is continuous and $f'(x)$ is sectionally continuous on the interval $[0,2\pi]$, and the definition of $f(x)$ is extended so that $f(x+2\pi) = f(x)$ for all real values of x, then $f(x)$ has a Fourier series of the form

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$$

which converges to $f(x)$ for all real values of x.

As for the second question, it is easy to show in general that $f(x)$ is the sum of the series, provided that the system Φ is complete. Since we have assumed that each $\phi_n(x)$ is continuous on the interval $[a,b]$, the partial sums of the Fourier series are continuous there as well. If the Fourier series converges uniformly, then it converges to some continuous function in $\mathcal{R}^2[a,b]$. Consider the continuous function

$$\delta(x) = \sum_{n=1}^{\infty} \langle f, \phi_n\rangle \, \phi_n(x) - f(x).$$

If the series converges uniformly, then it can be integrated term by term to obtain $\langle \delta, \phi_n\rangle = 0$ for every $n \geq 1$. But since Φ is complete, it must be the case that $\|\delta\|_2 = 0$, i.e., that $\delta(x) \equiv 0$ on the interval $[a,b]$.

3.2 Hermitian Kernels and Operators

Let $K(x,t)$ be a complex-valued, continuous kernel defined on the square $Q(a,b)$. If

$$K^*(x,t) = \overline{K(t,x)} = K(x,t),$$

on $Q(a,b)$, then $K(x,t)$ is called a *Hermitian kernel*. If $K(x,t)$ is complex-valued and $K(x,t) = K(t,x)$, then $K(x,t)$ is referred to as a *complex symmetric kernel*. If $K(x,t)$ is real-valued and $K(x,t) = K(t,x)$, then $K(x,t)$ is called a *real symmetric kernel*. Every real symmetric kernel is Hermitian, but nonzero complex symmetric kernels are not.

Illustrative Examples

- *Example 1*: The kernels $A(x,t) = x+t$ and $B(x,t) = \mathrm{e}^{xt}$ are real symmetric, but $C(x,t) = x-t$ is not.
- *Example 2*: The kernels $D(x,t) = \mathrm{i}(x-t)$ and $E(x,t) = \mathrm{e}^{\mathrm{i}(x-t)}$ are Hermitian, but $F(x,t) = \mathrm{i}(x+t)$ is not. However, $F(x,t)$ is complex symmetric, since $F(x,t) = F(t,x)$.
- *Example 3*: If $G(x,t)$ is an arbitrary real-valued continuous kernel defined on the square $Q(a,b)$, then the composed kernels

  $$G_L(x,t) = \int_a^b G(x,s)\,G(t,s)\,\mathrm{d}s$$

  and

  $$G_R(x,t) = \int_a^b G(s,x)\,G(s,t)\,\mathrm{d}s$$

  are real symmetric kernels. If $G(x,t)$ is symmetric, then $G_L(x,t) = G_R(x,t)$.
- *Example 4*: If $\{\psi_i(x)\}_{i=1}^n$ is an arbitrarily chosen linearly independent set of complex-valued continuous functions defined on $[a,b]$ and each μ_i is real, then

  $$H(x,t) = \sum_{i=1}^{n} \frac{\psi_i(x)\,\overline{\psi_i(t)}}{\mu_i}$$

  is a Hermitian kernel. If, in addition, $\|\psi_i\|_2 = 1$, then $\phi(x) = \psi_i(x)$ is a solution to the homogeneous integral equation

  $$\phi(x) = \mu_i \int_a^b H(x,t)\,\phi(t)\,\mathrm{d}t$$

for each $i = 1, \ldots, n$. It follows that each μ_i is an eigenvalue and each $\phi_i(x)$ is a normalized eigenfunction of the kernel $H(x,t)$. Kernels of this form will play a significant role in the theory to follow.

The Fredholm operator $\mathbf{K}$ is called a *Hermitian operator* if $\mathbf{K}^* = \mathbf{K}$. Any Hermitian operator enjoys the property

$$\langle \mathbf{K}\phi, \psi \rangle = \langle \phi, \mathbf{K}\psi \rangle$$

for all $\phi, \psi \in \mathcal{R}^2[a,b]$.

If $K(x,t)$ is a Hermitian kernel, then $\mathbf{K}$ is a Hermitian operator, since

$$\begin{aligned} \mathbf{K}\phi &= \int_a^b K(x,t)\,\phi(t)\,\mathrm{d}t \\ &= \int_a^b K^*(x,t)\,\phi(t)\,\mathrm{d}t \\ &= \mathbf{K}^*\phi \end{aligned}$$

for every $\phi \in \mathcal{R}^2[a,b]$.

If $\mathbf{K}$ is any Fredholm operator, then the composition $\mathbf{K}\mathbf{K}^*$ is a Hermitian operator, since $(\mathbf{K}\mathbf{K}^*)^* = \mathbf{K}^{**}\mathbf{K}^* = \mathbf{K}\mathbf{K}^*$. Similarly, $\mathbf{K}^*\mathbf{K}$ is Hermitian.

The Fredholm operator $\mathbf{K}$ is called a *normal operator* if $\mathbf{K}\mathbf{K}^* = \mathbf{K}^*\mathbf{K}$. If $\mathbf{K}$ is a normal operator, then

$$\begin{aligned} \|\mathbf{K}\phi\|_2^2 &= \langle \mathbf{K}\phi, \mathbf{K}\phi \rangle \\ &= \langle \phi, \mathbf{K}^*\mathbf{K}\phi \rangle \\ &= \langle \phi, \mathbf{K}\mathbf{K}^*\phi \rangle \\ &= \langle \phi, \mathbf{K}^{**}\mathbf{K}^*\phi \rangle \\ &= \langle \mathbf{K}^*\phi, \mathbf{K}^*\phi \rangle \\ &= \|\mathbf{K}^*\phi\|_2^2 \end{aligned}$$

for all $\phi \in \mathcal{R}^2[a,b]$.

A linear operator $\mathbf{K}\colon \mathcal{R}^2[a,b] \to \mathcal{R}^2[a,b]$ is called *continuous* if there exists a constant C such that

$$\|\mathbf{K}f - \mathbf{K}g\|_2 \le C\,\|f - g\|_2.$$

If $\mathbf{K}$ and $\mathbf{L}$ are two continuous operators, then their sum and difference are also continuous. In addition, if c is any constant, then $c\mathbf{K}$ is continuous.

If $K(x,t)$ is any complex-valued, continuous, Hermitian kernel, then the corresponding Fredholm operators $\mathbf{K}$ and $\mathbf{K}^*$ are continuous. Since we always have $\|\mathbf{K}\| \le \|K\|_2 < +\infty$, we can choose $C = \|\mathbf{K}\|$.

If $\{\phi_n\}$ is a strongly convergent sequence and $\mathbf{K}$ is any continuous linear operator, then $\{\mathbf{K}\phi_n\}$ is also a strongly convergent sequence, since

$$\|\mathbf{K}\phi - \mathbf{K}\phi_n\|_2 \leq \|\mathbf{K}\|\,\|\phi - \phi_n\|_2.$$

A sequence $\{\phi_n\}$ in $\mathcal{R}^2(a,b)$ is said to be *weakly convergent to an element* ϕ if $\langle \phi_n, \psi\rangle \to \langle \phi, \psi\rangle$ as $n \to +\infty$ for all $\psi \in \mathcal{R}^2(a,b)$. An operator $\mathbf{K}$ is called *completely continuous* if $\{\mathbf{K}\phi_n\}$ converges strongly to $\mathbf{K}\phi$ whenever $\{\phi_n\}$ converges weakly to ϕ.

3.2.1 Properties of Hermitian Kernels and Operators

Hermitian kernels have many properties that are interesting in their own right.

Proposition 3.2.1 (Iterated Kernels). *Suppose that $K(x,t)$ is a continuous, complex-valued, Hermitian kernel defined on the square $Q(a,b)$.*

Then, for $m \geq 1$, the iterated kernels $K_m(x,t)$ satisfy each of the following properties:

1. *Each $K_m(x,t)$ is complex-valued, continuous, and Hermitian.*
2. *$\|K_m\|_2 \leq \|K\|_2^m < +\infty$.*
3. *Each $K_m(x,x)$ is real-valued on the interval $[a,b]$.*
4. *If $K(x,t) \not\equiv 0$ on $Q(a,b)$, then each iterated kernel $K_m(x,t) \not\equiv 0$ there.*
5. *If $\mathbf{K}^m$ denotes the Fredholm operator corresponding to $K_m(x,t)$, then the inner products $\langle \mathbf{K}^m\phi, \phi\rangle$ are real-valued for every $\phi \in \mathcal{R}^2[a,b]$.*

Proof. 1. The proof proceeds by induction. The kernel $K_2(x,t)$ is Hermitian since

$$\begin{aligned} K_2(x,t) &= \int_a^b K(x,s)\,K(s,t)\,\mathrm{d}s \\ &= \int_a^b \overline{K(t,s)}\,\overline{K(s,x)}\,\mathrm{d}s \\ &= \overline{K_2(t,x)} \\ &= K_2^*(x,t). \end{aligned}$$

If $K_m(x,t)$ is Hermitian, then

$$\begin{aligned} K_{m+1}(x,t) &= \int_a^b K(x,s)\,K_m(s,t)\,\mathrm{d}s \\ &= \int_a^b \overline{K_m(t,s)}\,\overline{K(s,x)}\,\mathrm{d}s \\ &= \overline{K_{m+1}(t,x)} \\ &= K_{m+1}^*(x,t). \end{aligned}$$

2. An application of the Cauchy–Schwarz inequality gives $\|K_2\|_2 \leq \|K\|_2^2$. The assertion follows by induction.
3. Since $K_m(x,t)$ is Hermitian, it follows that $K_m(x,x) = \overline{K_m(x,x)}$ for every $x \in [a,b]$.
4. The proof proceeds by contradiction. Suppose that $K_m(x,t) \equiv 0$ on the square $Q(a,b)$ and that $m > 1$ is the smallest integer for which this is the case. Then $K_n(x,t) \not\equiv 0$ for $n = 1,\ldots,m-1$ and $K_n(x,t) \equiv 0$ if $n \geq m$. Let $2N = m$ or $m+1$, whichever is even. An application of the iteration formula for kernels yields

$$K_{2N}(x,t) = \int_a^b K_N(x,s)\,K_N(s,t)\,\mathrm{d}s \equiv 0$$

on the square $Q(a,b)$. In particular, if $x \in [a,b]$, then

$$\begin{aligned} K_{2N}(x,x) &= \int_a^b K_N(x,s)\,K_N(s,x)\,\mathrm{d}s \\ &= \int_a^b K_N(x,s)\,\overline{K_N(x,s)}\,\mathrm{d}s \\ &= \int_a^b |K_N(x,s)|^2\,\mathrm{d}s \equiv 0, \end{aligned}$$

so that $K_N(x,s) \equiv 0$ as a function of s for each $x \in [a,b]$, i.e., $K_N(x,t) \equiv 0$ on the square $Q(a,b)$. But this conclusion contradicts the assumption that m was the smallest integer for which $K_m(x,t) \equiv 0$, since $N < m$.
5. The assertion follows directly from the relations

$$\langle \mathbf{K}^m\phi,\phi\rangle = \langle \phi,\mathbf{K}^{m*}\phi\rangle = \langle \phi,\mathbf{K}^m\phi\rangle = \overline{\langle \mathbf{K}^m\phi,\phi\rangle}. \qquad \square$$

3.2.2 The Eigenvalues of a Hermitian Kernel

To show that nonvanishing, continuous, Hermitian kernels have eigenvalues, it is necessary to establish beforehand the following result regarding the trace A_m of the iterated kernel $K_m(x,t)$, or what is the same, the mth trace of the kernel $K(x,t)$. Recall from Sect. 2.5 that

$$A_m = \int_a^b K_m(x,x)\,\mathrm{d}x.$$

Lemma 3.2.1 (The Trace Lemma). *Let $K(x,t)$ be a nonvanishing, continuous, Hermitian kernel.*

Then the traces A_m satisfy the following properties:

1. *For every $m \geq 1$, A_m is a real number.*
2. *For every $m \geq 1$, $A_{2m} = \|K_m\|_2^2 > 0$.*
3. *If k and p are positive integers, then $A_{k+p}^2 \leq A_{2k}A_{2p}$.*

4. *The sequence* $\left\{\dfrac{A_{2n+2}}{A_{2n}}\right\}$ *is nondecreasing and bounded below by* A_4/A_2. *In particular,*

$$0 < \frac{A_4}{A_2} \leq \limsup_{n\to\infty} \frac{A_{2m+2}}{A_{2m}}.$$

Proof. The proof of this lemma follows directly from Proposition 3.2.1.

1. The values $K_m(x,x)$ are real for every $m \geq 1$ and every $x \in [a,b]$.
2. Since $K_m(x,t)$ is Hermitian for every $m \geq 1$, we have

$$\begin{aligned} A_{2m} &= \int_a^b K_{2m}(x,x)\,\mathrm{d}x \\ &= \int_a^b \int_a^b K_m(x,t)\,K_m(t,x)\,\mathrm{d}t\,\mathrm{d}x \\ &= \int_a^b \int_a^b K_m(x,t)\,\overline{K_m(x,t)}\,\mathrm{d}t\,\mathrm{d}x \\ &= \int_a^b \int_a^b |K_m(x,t)|^2\,\mathrm{d}t\,\mathrm{d}x \\ &= \|K_m\|^2. \end{aligned}$$

   Since $K_m(x,t) \not\equiv 0$, its norm is positive.
3. Since $K(x,t)$ is Hermitian, the iteration formula for kernels assumes the form

$$K_{k+p}(x,x) = \int_a^b K_k(x,t)\,\overline{K_p(x,t)}\,\mathrm{d}t.$$

   Since $\|K_m\|_2 < +\infty$ for all $m \geq 1$, an application of the Cauchy–Schwarz Inequality yields

$$\begin{aligned} A_{k+p}^2 &= \left(\int_a^b K_{k+p}(x,x)\,\mathrm{d}x\right)^2 \\ &= \left(\int_a^b \int_a^b K_k(x,t)\,\overline{K_p(x,t)}\,\mathrm{d}t\,\mathrm{d}x\right)^2 \\ &\leq \left(\int_a^b \int_a^b |K_k(x,t)|^2\,\mathrm{d}x\,\mathrm{d}t\right)\left(\int_a^b \int_a^b |K_p(x,t)|^2\,\mathrm{d}x\,\mathrm{d}t\right) \\ &= \left(\int_a^b K_{2k}(x,x)\,\mathrm{d}x\right)\left(\int_a^b K_{2p}(x,x)\,\mathrm{d}x\right) \\ &= A_{2k}A_{2p}. \end{aligned}$$

4. If we set $k = n-1$ and $p = n+1$, then as a special case, we obtain

$$A_{2n}^2 \leq A_{2n-2}A_{2n+2}.$$

Since the traces with even subscripts are positive, we can rearrange this inequality to obtain the sequence of inequalities

$$\frac{A_4}{A_2} \leq \frac{A_6}{A_4} \leq \cdots \leq \frac{A_{2n}}{A_{2n-2}} \leq \frac{A_{2n+2}}{A_{2n}} \leq \cdots \qquad \square$$

With the Trace Lemma, we can establish the most important property of Hermitian kernels.

Proposition 3.2.2. *If $K(x,t)$ is a nonvanishing, continuous, Hermitian kernel, then $K(x,t)$ and its iterates $K_m(x,t)$ have at least one eigenvalue. Furthermore, if λ_1 is the eigenvalue of smallest modulus of $K(x,t)$, then $|\lambda_1| \leq \sqrt{A_2/A_4}$.*

Proof. The proof proceeds by contradiction. Suppose that $K(x,t)$ does not have any eigenvalues. Then, for each $x,t \in [a,b]$, the resolvent kernel defined in the Theorem of Successive Approximation in Sect. 2.3 is an entire function of λ that can be represented by the series

$$R(x,t;\lambda) = \sum_{m=1}^{\infty} K_m(x,t)\,\lambda^{m-1}$$

which converges for every complex value of λ. If we set $t = x$ and integrate the result, then we obtain the series

$$\lambda \int_a^b R(x,x;\lambda)\,\mathrm{d}x = \int_a^b \left(\sum_{m=1}^{\infty} K_m(x,x)\,\lambda^m \right) \mathrm{d}x = \sum_{m=1}^{\infty} A_m \lambda^m,$$

which also converges for every complex λ. The even subseries

$$\sum_{m=1}^{\infty} A_{2m} |\lambda|^{2m} \tag{3.1}$$

also has an infinite radius of convergence. By the Trace Lemma, we have

$$0 < \frac{A_4}{A_2} |\lambda|^2 \leq \limsup_{m\to\infty} \frac{A_{2m+2}}{A_{2m}} |\lambda|^2 = \rho.$$

If $|\lambda| > \sqrt{A_2/A_4}$, then $\rho > 1$. But if $\rho > 1$, then the series (3.1) diverges by Cauchy's ratio test. Hence, the assumption that $K(x,t)$ had no eigenvalues was false. Also, the required inequality follows from the fact that $\rho \leq 1$. The iterates $K_m(x,t)$ have at least one eigenvalue as well by Proposition 3.2.1(1, 4). $\square$

Proposition 3.2.3. *If $K(x,t)$ is a complex-valued, nonvanishing, continuous, Hermitian kernel, then the eigenvalues of $K_m(x,t)$ are real for all $m \geq 1$.*

Proof. If $\phi = \mu\,\mathbf{K}^m\phi$, then $\|\phi\|^2 = \mu\,\langle \mathbf{K}^m\phi, \phi\rangle$. Therefore, μ is real since $\langle \mathbf{K}^m\phi, \phi\rangle$ is real by Proposition 3.2.1(5). $\square$

Note: Even though Hermitian kernels have real eigenvalues, they may have complex-valued eigenfunctions. Consider the integral equation

$$\phi(x) = \lambda \int_0^{2\pi} \mathrm{e}^{\mathrm{i}(x-t)}\, \phi(t)\, \mathrm{d}t.$$

The only eigenvalue is $\lambda_1 = 1/2\pi$, and the corresponding eigenfunction is $\phi_1(x) = \mathrm{e}^{\mathrm{i}x}/\sqrt{2\pi}$.

The next result will be used in the proof of the Hilbert–Schmidt Theorem given in Sect. 3.4, and will also serve as a theoretical basis for the computation of eigenvalues in Sect. 3.5. We motivate it with the following observation.

If λ_n is any eigenvalue of the Hermitian kernel $K(x,t)$ and ϕ_n is one of its corresponding normalized eigenfunctions, then $\phi_n = \lambda_n \mathbf{K}\phi_n$. Forming the inner product with ϕ_n, we obtain

$$1 = \|\phi_n\|_2^2 = \langle \phi_n, \phi_n \rangle = \lambda_n \langle \mathbf{K}\phi_n, \phi_n \rangle.$$

After division, we have

$$\frac{1}{|\lambda_n|} = |\langle \mathbf{K}\phi_n, \phi_n \rangle|.$$

If the eigenvalues of $K(x,t)$ are arranged in order of increasing magnitude, i.e., $|\lambda_1| \le |\lambda_2| \le \cdots$, then

$$\frac{1}{|\lambda_1|} = \max_n |\langle \mathbf{K}\phi_n, \phi_n \rangle|.$$

A stronger result can be proven.

Theorem 3.2.1. *Let $K(x,t)$ be a nonvanishing, continuous, Hermitian kernel. If λ_1 is the eigenvalue of smallest modulus of $K(x,t)$ and $\mathbf{K}$ is the corresponding Hermitian operator, then*

$$\frac{1}{|\lambda_1|} = \max_{\phi} |\langle \mathbf{K}\phi, \phi \rangle|,$$

where the maximum is taken over all $\phi \in \mathcal{R}^2[a,b]$ subject to the condition that

$$\|\phi\|^2 = \langle \phi, \phi \rangle = \int_a^b |\phi(x)|^2\, \mathrm{d}x = 1.$$

Furthermore, this maximum value is attained when $\phi(x)$ is an eigenfunction of the kernel corresponding to λ_1.

Proof. Let $K(x,t)$ be a nonvanishing, continuous, Hermitian kernel, and let $\mathbf{K}$ be the corresponding Fredholm operator with necessarily finite norm $\|\mathbf{K}\|$. For any $\phi \in \mathcal{R}^2[a,b]$ with $\|\phi\|_2 = 1$, we have

$$|\langle \mathbf{K}\phi, \phi \rangle| \le \|\mathbf{K}\phi\|_2 \cdot \|\phi\|_2 \le \|\mathbf{K}\| \cdot \|\phi\|_2^2 = \|\mathbf{K}\|$$

as a direct consequence of the Cauchy–Schwarz inequality.

Let B denote the least upper bound of the bounded set

$$\mathcal{K} = \{|\langle \mathbf{K}\phi, \phi\rangle| \colon \|\phi\|_2 = 1\}.$$

By the definition of the least upper bound, there exists at least one sequence $\{\psi_n(x)\}_{n=1}^{\infty}$ in $\mathcal{R}^2[a,b]$ with $\|\psi_n\|_2 = 1$ such that

$$\lim_{n\to\infty} |\langle \mathbf{K}\psi_n, \psi_n\rangle| = B.$$

If the sequence $\langle \mathbf{K}\psi_n, \psi_n\rangle$ of real numbers converges, then it must converge to either $+B$ or $-B$. If it does not converge, then there must be at least two subsequences, one of which converges to $+B$ and another one which converges to $-B$. Without loss of generality, it suffices to assume that

$$\lim_{n\to\infty} \langle \mathbf{K}\psi_n, \psi_n\rangle = +B.$$

Since the sequence $\{\psi_n(x)\}_{n=1}^{\infty}$ is bounded and $\|K\|_2$ is finite, the sequence $\{\mathbf{K}\psi_n(x)\}_{n=1}^{\infty}$ is a compact set in $\mathcal{R}^2[a,b]$. Thus, there must exist a function $w \in \mathcal{R}^2[a,b]$ such that

$$w(x) = \lim_{n\to\infty} \mathbf{K}\psi_n.$$

Set $\phi_1(x) = w(x)/B$ and $\lambda_1 = 1/B$. If we can show that

$$\lim_{n\to\infty} \psi_n(x) = \frac{w(x)}{B} \quad \text{and} \quad \lim_{n\to\infty} \mathbf{K}\psi_n = \mathbf{K}\phi_1,$$

it will follow that

$$\phi_1(x) = \frac{w(x)}{B} = \lim_{n\to\infty} \psi_n(x) = \frac{1}{B} \lim_{n\to\infty} \mathbf{K}\psi_n = \lambda_1 \mathbf{K}\phi_1(x).$$

Also, $\|\phi_1\|_2 = 1$, since $\|\psi_n\|_2 = 1$ for all n.

A clever variational argument allows us to complete the proof. For any real t and any $\alpha_n \in \mathcal{R}^2[a,b]$, we can construct the competing functions

$$\beta_n(x,t) = \frac{\psi_n(x) + t\,\alpha_n(x)}{\|\psi_n(x) + t\,\alpha_n(x)\|_2} \in \mathcal{R}^2[a,b]$$

with $\|\beta_n\|_2 = 1$ for every $n \geq 1$. Since the real number $|\langle \mathbf{K}\beta_n, \beta_n\rangle| \in \mathcal{K}$, the inequality $\langle \mathbf{K}\beta_n, \beta_n\rangle \leq B$ holds for all $n \geq 1$. After expanding this inequality in powers of t, we obtain

$$(\langle \mathbf{K}\phi_n, \phi_n\rangle - B) + (\mathrm{Re}\{\langle \mathbf{K}\psi_n - B\,\psi_n, \alpha_n\rangle\})\,t + \left(\langle \mathbf{K}\alpha_n, \alpha_n\rangle - B\,\|\alpha_n\|^2\right) t^2 \leq 0$$

for every $n \geq 1$. The discriminant of this quadratic polynomial cannot be positive. That is,

$$|\mathrm{Re}\{\langle \mathbf{K}\psi_n - B\,\psi_n, \alpha_n\rangle\}| \leq \sqrt{B\|\alpha_n\|^2 - \langle \mathbf{K}\alpha_n, \alpha_n\rangle}\,\sqrt{B - \langle \mathbf{K}\psi_n, \psi_n\rangle}$$

for every $n \geq 1$ and every choice of α_n. If we choose $\alpha_n = \mathbf{K}\psi_n - B\,\psi_n$, then the first term on the right in the discriminant inequality is bounded by $(B + \|\mathbf{K}\|)^{3/2}$. Since

$$\|\alpha_n\|_2 = \|\mathbf{K}\psi_n - B\,\psi_n\|_2 \leq B + \|\mathbf{K}\phi_n\|_2 \leq B + \|\mathbf{K}\|$$

and

$$|\langle \mathbf{K}\alpha_n, \alpha_n\rangle| \leq \|\mathbf{K}\alpha_n\|_2 \cdot \|\alpha_n\|_2 \leq \|\mathbf{K}\| \cdot \|\alpha_n\|_2^2,$$

we have

$$|B\|\alpha_n\|_2^2 - \langle \mathbf{K}\alpha_n, \alpha_n\rangle| \leq B\,(B + \|\mathbf{K}\|)^2 + \|\mathbf{K}\|\,(B + \|\mathbf{K}\|)^2 = (B + \|\mathbf{K}\|)^3.$$

Since the second term on the right in the discriminant inequality tends to 0, it follows that

$$\lim_{n\to\infty} \mathrm{Re}\,\langle \mathbf{K}\phi_n - B\,\phi_n, \mathbf{K}\phi_n - B\,\phi_n\rangle = \lim_{n\to\infty} \|\mathbf{K}\phi_n - B\,\phi_n\|^2 = 0,$$

from which we conclude that

$$\lim_{n\to\infty} \phi_n(x) = \lim_{n\to\infty} \frac{1}{B}\,\mathbf{K}\phi_n = \frac{1}{B}\,w(x) = \phi_1(x).$$

Also, since $\mathbf{K}$ is a continuous Fredholm operator, we know that

$$\begin{aligned}\|\mathbf{K}\phi_n - \mathbf{K}\phi_1\|_2 &= \left\|\mathbf{K}\left(\phi_n(x) - \frac{w(x)}{B}\right)\right\|_2 \\ &\leq \|\mathbf{K}\| \cdot \left\|\phi_n(x) - \frac{w(x)}{B}\right\|_2 \to 0\end{aligned}$$

as $n \to \infty$, so that

$$\lim_{n\to\infty} \frac{1}{B}\,\mathbf{K}\phi_n = \frac{1}{B}\,\mathbf{K}\phi_1 = \lambda_1\,\mathbf{K}\phi_1.$$

Hence, $\phi_1 = \lambda_1 K\phi_1$, and $\phi_1(x) \not\equiv 0$, since $\|\phi_1\|_2 = 1$. □

Proposition 3.2.4. *Let $K(x,t)$ be a kernel (whether Hermitian or not) for which $\|K\|_2 < +\infty$. Let $\Lambda = \{\lambda_i\}$ denote the set of all eigenvalues of the kernel $K(x,t)$ and, for $m \geq 2$, let $\Lambda^m = \{\lambda_i^m\}$ denote the set of all mth powers of these eigenvalues. If Λ_m denotes the set of all eigenvalues of the iterated kernel $K_m(x,t)$, then $\Lambda^m = \Lambda_m$.*

Proof. ($\Lambda^m \subseteqq \Lambda_m$) If $\lambda \in \Lambda$, then

$$\phi(t) = \lambda \int_a^b K(t,s)\,\phi(s)\,\mathrm{d}s$$

for some corresponding nontrivial eigenfunction $\phi(t)$. Multiply this equation by $\lambda K(x,t)$ and then integrate the result from a to b to obtain

$$\phi(x) = \lambda \int_a^b K(x,t)\,\phi(t)\,\mathrm{d}t = \lambda^2 \int_a^b K_2(x,s)\,\phi(s)\,\mathrm{d}s.$$

This shows that λ^2 is an eigenvalue of the iterated kernel $K_2(x,t)$ with the same corresponding eigenfunction. By continuing this process, we obtain

$$\phi(x) = \lambda^m \int_a^b K_m(x,s)\,\phi(s)\,\mathrm{d}s$$

for every integer $m \geq 2$, showing that λ^m is an eigenvalue of the iterated kernel $K_m(x,t)$ with the same corresponding eigenfunction. Hence, $\lambda^m \in \Lambda_m$.

($\Lambda^m \supseteq \Lambda_m$) If $\rho \in \Lambda_m$, then there exists a corresponding eigenfunction $\psi(x) \not\equiv 0$ such that

$$\psi(x) = \rho \int_a^b K_m(x,t)\,\psi(t)\,\mathrm{d}t.$$

We must show that there exists at least one eigenvalue $\lambda \in \Lambda$ for which $\lambda^m = \rho$. Let $\{e_1,\ldots,e_m\}$ be the m distinct mth roots of ρ, and define the functions $E_i(x)$ by setting

$$mE_i(x) = \psi(x) + \sum_{j=1}^{m-1} e_i^j \int_a^b K_j(x,t)\,\psi(t)\,\mathrm{d}t.$$

If we multiply both sides of this definition by $e_i K(y,x)$ and integrate the result with respect to x, then we obtain

$$\begin{aligned}
me_i \int_a^b K(y,x)\,E_i(x)\,\mathrm{d}x &= e_i \int_a^b K(y,x)\,\psi(x)\,\mathrm{d}x \\
&\quad + \sum_{j=1}^{m-1} e_i^{j+1} \int_a^b K_{j+1}(y,t)\,\psi(t)\,\mathrm{d}t \\
&= e_i \int_a^b K(y,x)\,\psi(x)\,\mathrm{d}x \\
&\quad + \sum_{j=2}^{m} e_i^j \int_a^b K_j(y,t)\,\psi(t)\,\mathrm{d}t \\
&= \sum_{j=1}^{m-1} e_i^j \int_a^b K_j(y,t)\,\psi(t)\,\mathrm{d}t \\
&\quad + \rho \int_a^b K_m(x,t)\,\psi(t)\,\mathrm{d}t \\
&= \psi(y) + \sum_{j=1}^{m-1} e_i^j \int_a^b K_j(y,t)\,\psi(t)\,\mathrm{d}t \\
&= mE_i(y).
\end{aligned}$$

To complete the proof, we must show that $E_i(x) \not\equiv 0$ for at least one index i. By adding all m of the equations in the definition and canceling a factor of m, we obtain

$$E_1(x) + \cdots + E_m(x) = \psi(x),$$

after using the well-known fact that the sum of the mth roots of unity satisfies the property

$$e_1^j + \cdots + e_m^j = 0$$

for every $j = 1, \ldots, m-1$. This equation implies that at least one of the functions $E_i(x) \not\equiv 0$, since $\psi(x) \not\equiv 0$. Choose $\lambda = e_i$. □

3.2.3 The Eigenfunctions of a Hermitian Kernel

Proposition 3.2.5. *Let $K(x,t)$ be a nonzero, continuous, Hermitian kernel. Then the set of eigenfunctions corresponding to any given eigenvalue λ_i can be orthonormalized.*

Proof. Since the set of eigenfunctions corresponding to any eigenvalue is linearly independent, the Gram–Schmidt Orthonormalization Procedure can be applied to convert it into an orthonormal set. □

Proposition 3.2.6. *Let $K(x,t)$ be a nonzero, continuous, Hermitian kernel. If λ_i and λ_j are any two different eigenvalues, then the eigenfunctions ϕ_i and ϕ_j that correspond to them are orthogonal.*

Proof. Suppose that $\phi_i = \lambda_i \mathbf{K}\phi_i$ and $\phi_j = \lambda_j \mathbf{K}\phi_j$. Then,

$$\frac{1}{\lambda_i}\langle \phi_i, \phi_j \rangle = \langle \mathbf{K}\phi_i, \phi_j \rangle = \langle \phi_i, \mathbf{K}^*\phi_j \rangle = \langle \phi_i, \mathbf{K}\phi_j \rangle = \frac{1}{\lambda_j}\langle \phi_i, \phi_j \rangle.$$

Since $\lambda_i \neq \lambda_j$, it must be the case that $\langle \phi_i, \phi_j \rangle = 0$. □

Proposition 3.2.7. *Let $K(x,t)$ be a nonzero, continuous, Hermitian kernel. Then the set of all of its eigenfunctions is a mutually orthonormal system of functions.*

Proof. The eigenfunctions that correspond to any given eigenvalue λ_i can be orthonormalized, and the eigenfunctions of unit norm corresponding to different eigenvalues are orthogonal. Thus, the set of all eigenfunctions of a Hermitian kernel constitutes a mutually orthonormal system of functions. □

Proposition 3.2.8. *Let $K(x,t)$ be a nonzero, continuous, Hermitian kernel. Then the full set of its eigenvalues can be enumerated in order of increasing magnitude. Consequently, the eigenfunctions corresponding to them can be listed in a sequence, although this listing need not always be unique.*

Proof. The Fourth Fredholm Theorem implies that there exist only a finite number of eigenvalues in any finite disk centered at the origin. In particular, there can only be a finite number of eigenvalues on any circle centered at the origin. If two eigenvalues have the same modulus, then they can be numbered according to their argument. Since there can only be a countably infinite number of eigenvalues of a Hermitian kernel, we can now write

$$|\lambda_1| \leq |\lambda_2| \leq \cdots \leq |\lambda_n| \leq \cdots .$$

Once the eigenvalues have been thus enumerated, the eigenfunctions can be enumerated correspondingly. □

For the remainder of this chapter, we will assume that the eigenvalues and eigenfunctions are so enumerated.

Illustrative Examples

- *Example 1*: A kernel $K(x,t)$ is *skew Hermitian* if $K^*(x,t) = -K(x,t)$. The eigenvalues of a skew-Hermitian kernel are purely imaginary. On the one hand, we have

  $$\langle K^*\phi, \phi\rangle = \langle \phi, K\phi\rangle = \left\langle \phi, \frac{\phi}{\lambda} \right\rangle = \frac{1}{\overline{\lambda}} \langle \phi, \phi\rangle.$$

  On the other hand, we also have

  $$\langle -K\phi, \phi\rangle = \left\langle -\frac{\phi}{\lambda}, \phi \right\rangle = -\frac{1}{\lambda} \langle \phi, \phi\rangle.$$

  Thus, $-\lambda = \overline{\lambda}$, or $\lambda + \overline{\lambda} = 2\,\mathrm{Re}\{\lambda\} = 0$.
- *Example 2*: The eigenvalues of the composed nonvanishing, continuous, and Hermitian kernels KK^* and K^*K are positive. For if $\phi = \lambda\, \mathbf{K}\mathbf{K}^*\phi$, then

  $$\|\phi\|^2 = \langle \lambda\, \mathbf{K}\mathbf{K}^*\phi, \phi\rangle = \lambda\, \langle \mathbf{K}^*\phi, \mathbf{K}^*\phi\rangle = \lambda\, \|\mathbf{K}^*\phi\|^2.$$

  The second assertion is established in a similar manner.
- *Example 3*: If $K(x,t)$ is a real symmetric kernel, then its iterates $K_m(x,t)$ and the corresponding resolvent $R(x,t;\lambda)$ are also real symmetric kernels. For if $K(x,t)$ is real symmetric, then

  $$K_2(x,t) = \int_a^b K(x,s)\, K(s,t)\, \mathrm{d}t = \int_a^b K(t,s)\, K(s,x)\, \mathrm{d}s = K_2(t,x),$$

  and the result follows for all iterates by a simple induction argument. Hence, we have

  $$R(x,t;\lambda) = \sum_{m=1}^{\infty} \lambda^{m-1} K_m(x,t) = \sum_{m=1}^{\infty} \lambda^{m-1} K_m(t,x) = R(t,x;\lambda).$$

- *Example 4*: Consider the integral equation

$$\phi(x) = f(x) + \lambda \int_a^b r(t)\, L(x,t)\, \phi(t)\, \mathrm{d}t,$$

where $r(t)$ is nonnegative and $L(x,t)$ is a real symmetric kernel. Then the kernel $r(t)\,L(x,t)$ in this equation is not symmetric. However, if we multiply this integral equation by $\sqrt{r(x)}$, then it becomes

$$\left(\sqrt{r(x)}\,\phi(x)\right) = f(x) + \lambda \int_a^b K(x,t)\, \left(\sqrt{r(t)}\phi(t)\right) \mathrm{d}t$$

where now the new kernel $K(x,t) = \sqrt{r(x)\,r(t)}\,L(x,t)$ is real symmetric.

- *Example 5*: If the kernel $K(x,t)$ is real symmetric, then its eigenfunctions may be assumed to be real. For if $\phi(x) = \alpha(x) + \mathrm{i}\,\beta(x)$ with $\alpha(x)$ and $\beta(x)$ real, then the two equations

$$\alpha(x) \pm \mathrm{i}\,\beta(x) = \lambda \int_a^b K(x,t)\, (\alpha(t) \pm \mathrm{i}\,\beta(t))\, \mathrm{d}t$$

can be added and subtracted to show that both the real part $\alpha(x)$ and the imaginary part $\beta(x)$ are eigenfunctions.

- *Example 6*: The inner product $\langle \mathbf{K}\phi, \phi \rangle$ can be written as a double integral. Indeed, we have

$$\begin{aligned} \langle \mathbf{K}\phi, \phi \rangle &= \int_a^b \left(\int_a^b K(x,t)\, \phi(t)\, \mathrm{d}t \right) \overline{\phi(x)}\, \mathrm{d}x \\ &= \int_a^b \int_a^b K(x,t)\, \overline{\phi(x)}\, \phi(t)\, \mathrm{d}x\, \mathrm{d}t. \end{aligned}$$

- *Example 7*: Consider the continuous symmetric kernel

$$K(x,t) = \frac{1}{1+xt}$$

defined on the square $Q(0,1)$. The trace A_1 of $K(x,t)$ is easily computed as

$$A_1 = \int_0^1 \frac{1}{1+x^2}\, \mathrm{d}x = \arctan 1 = \frac{\pi}{4}.$$

To compute A_2, we first compute the iterated kernel $K_2(x,t)$ as

$$K_2(x,t) = \int_0^1 \frac{1}{1+xs}\, \frac{1}{1+st}\, \mathrm{d}s = \frac{\ln(1+x) - \ln(1+t)}{x-t}.$$

We recognize $K_2(x,t)$ to be the difference quotient for the derivative of $\ln(1+x)$. Hence,

$$A_2 = \int_0^1 K_2(x,x)\,\mathrm{d}x = \int_0^1 \frac{1}{1+x}\,\mathrm{d}x = \ln 2.$$

The trace A_2 can be computed in another way. For if we substitute the geometric series for the kernel into the integrand, we obtain

$$K_2(x,t) = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} \frac{(-1)^{m+n}}{m+n+1}\,x^n\,t^m$$

from which we conclude that

$$A_2 = \int_0^1 K_2(x,x)\,\mathrm{d}x = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} \frac{(-1)^{m+n}}{(m+n+1)^2} = \sum_{N=0}^{\infty} \frac{(-1)^N}{N+1} = \ln 2.$$

If it is required to compute A_3, then the series method would be easier.

Section 3.2 Exercises

1. Suppose that the kernel $K(x,t)$ is continuous in the variables x and t. Are the corresponding eigenfunctions $\phi_n(x)$ continuous?
   Hint: Consider the difference

$$\phi_n(x) - \phi_n(y) = \lambda_n \int_a^b (K(x,t) - K(y,t))\,\phi_n(t)\,\mathrm{d}t.$$

2. Consider the continuous symmetric kernel

$$K(x,t) = \frac{1}{1+x+t}$$

   defined on the square $Q(0,1)$. Show that $A_1 = \frac{1}{2}\ln 3$. Show that

$$K_2(x,t) = \frac{\ln(x+1) - \ln(t+1)}{x-t} - \frac{\ln(x+2) - \ln(t+2)}{x-t}$$

   and then use this fact to show that $A_2 = \ln\left(\frac{4}{3}\right)$.
3. Consider the continuous symmetric kernel

$$K(x,t) = \mathrm{e}^{xt}$$

defined on the square $Q(0,1)$. Show that

$$A_1 = \int_0^1 \mathrm{e}^{x^2}\,\mathrm{d}x = \sum_{n=0}^{\infty} \frac{1}{n!\,(2n+1)} \approx 1.462651.$$

Use the series representation for the exponential function to show that

$$K_2(x,t) = \sum_{n,m=0}^{\infty} \frac{1}{n!\,m!\,(n+m+1)}\, x^n t^m$$

and then conclude that

$$A_2 = \sum_{n,m=0}^{\infty} \frac{1}{n!\,m!\,(n+m+1)^2} \approx 1.841935.$$

In a similar fashion, show that

$$A_3 = \sum_{n,m,k=0}^{\infty} \frac{1}{n!\,m!\,k!\,(n+m+1)(m+k+1)(k+n+1)} \approx 2.478170.$$

In a similar manner, show that $A_4 \approx 3.351554$. Then use the inequality $|\lambda_1| \leq \sqrt{A_2/A_4}$ that was proven in Proposition 3.2.2 to estimate $|\lambda_1|$. Finally, derive a general formula for A_N.

4. Compute the eigenvalues and eigenvectors of the skew-symmetric kernel $K(x,t) = x\ln t - t\ln x$ defined on the square $Q(1,e)$. (Recall from Example 1 that the eigenvalues of a skew-symmetric kernel are purely imaginary.)
5. Suppose that all of the eigenvalues of some kernel are greater than one. Explain why the sequence $\{A_m\}_{m=1}^{\infty}$ of traces is strictly decreasing.
6. Suppose that all of the eigenvalues of some kernel are positive and that $A_1 < A_2$. What can you conclude from this information?
7. Suppose that all of the eigenvalues of some kernel are positive. Determine upper and lower bounds on the quantity $A_m^{1/m}$. Does the limit

$$\lim_{m\to+\infty} A_m^{1/m}$$

exist? If so, what is the value of this limit? If not, why not?

8. Many interesting real symmetric kernels are defined in terms of cases in the following manner:

$$K(x,t) = \begin{cases} g(x,t) & \text{if } 0 \leq x \leq t \leq 1 \\ g(t,x) & \text{if } 0 \leq t \leq x \leq 1. \end{cases}$$

For each of the kernels below, compute the iterated kernel $K_2(x,t)$, and the traces A_1 and A_2. (Recall that A_2 can be computed directly from $K_2(x,t)$ or from the formula $A_2 = \|K\|_2^2$ established in Lemma 3.2.1(2).)

(a) $K(x,t) = \begin{cases} x & \text{if } 0 \le x \le t \le 1 \\ t & \text{if } 0 \le t \le x \le 1 \end{cases}$

(b) $K(x,t) = \begin{cases} e^x & \text{if } 0 \le x \le t \le 1 \\ e^t & \text{if } 0 \le t \le x \le 1 \end{cases}$

(c) $K(x,t) = \begin{cases} x(1-t) & \text{if } 0 \le x \le t \le 1 \\ t(1-x) & \text{if } 0 \le t \le x \le 1 \end{cases}$

(d) $K(x,t) = \begin{cases} \sin x \cos t & \text{if } 0 \le x \le t \le 1 \\ \sin t \cos x & \text{if } 0 \le t \le x \le 1 \end{cases}$

Hint: If $0 \le x \le t \le 1$, then

$$K_2(x,t) = \left(\int_0^x + \int_x^t + \int_t^1 \right) K(x,s) K(s,t)\, \mathrm{d}s.$$

What are the values of $K(x,s)$ and $K(s,t)$ if s belongs to each of these three subintervals?

3.3 Finite Bilinear Expansions

The lemma below serves to illuminate the structure of Hermitian kernels with a finite number of eigenvalues. Not only is it interesting in its own right, but it also serves to motivate the discussion to follow in the rest of the chapter.

Lemma 3.3.1 (Kernel Structure). *Let $K(x,t)$ be a complex-valued, continuous, Hermitian kernel defined on the square $Q(a,b)$. Suppose that its eigenvalues $\lambda_1, \lambda_2, \ldots$ are enumerated in order of increasing magnitude with corresponding orthonormal eigenfunctions $\phi_1(x), \phi_2(x), \ldots$*

Then, the truncated kernel

$$\Delta_{N+1}(x,t) = K(x,t) - \sum_{n=1}^{N} \frac{\phi_n(x)\overline{\phi_n(t)}}{\lambda_n}$$

is also complex-valued, continuous, and Hermitian.

If $\Delta_{N+1}(x,t) \equiv 0$ for some $N \ge 1$ on $Q(a,b)$, then $K(x,t)$ assumes the canonical form of the finite bilinear expansion

$$K(x,t) = \sum_{n=1}^{N} \frac{\phi_n(x)\overline{\phi_n(t)}}{\lambda_n}.$$

In this case, the kernel $K(x,t)$ is separable, and it has exactly N eigenvalues counting multiplicity. The geometric multiplicity of each of these eigenvalues is equal to its algebraic multiplicity.

If $\Delta_{N+1}(x,t) \not\equiv 0$ for any $N \geq 1$on $Q(a,b)$, then the following statements are true:

1. *The eigenvalues $\lambda_1, \ldots, \lambda_N$ of $K(x,t)$ are not eigenvalues of $\Delta_{N+1}(x,t)$, and the corresponding eigenfunctions $\phi_1(x), \ldots, \phi_N(x)$ of $K(x,t)$ are not eigenfunctions of $\Delta_{N+1}(x,t)$.*
2. *The eigenvalues $\lambda_{N+1}, \lambda_{N+2}, \ldots$ of $K(x,t)$ are eigenvalues of $\Delta_{N+1}(x,t)$, and the corresponding eigenfunctions $\phi_{N+1}(x), \phi_{N+2}(x), \ldots$ of $K(x,t)$ are eigenfunctions of $\Delta_{N+1}(x,t)$.*
3. *The kernel $\Delta_{N+1}(x,t)$ has no eigenvalues or eigenfunctions other than those mentioned in the previous item.*

Proof. The truncated kernel is the difference of two complex-valued, continuous, Hermitian kernels.

Suppose that there exists an N such that $\Delta_{N+1}(x,t) \equiv 0$ on $Q(a,b)$. In this case, it is clear that $K(x,t)$ assumes the given separable form. Furthermore, the matrices $\mathbf{A} = (a_{ij}) = (\delta_{ij}/\lambda_j)$ and $\mathbf{I} - \lambda\,\mathbf{A}$ are diagonal, due to the assumed orthonormality of the eigenfunctions. Since the geometric multiplicities of the eigenvalues of the matrix $\mathbf{A}$ are equal to their algebraic multiplicities, the same is true for the eigenvalues of the kernel.

Now suppose that $\Delta_{N+1}(x,t) \not\equiv 0$ on $Q(a,b)$. In this case, we can prove that the truncation process has effectively removed the eigenvalues $\lambda_1, \ldots, \lambda_N$ and their corresponding eigenfunctions $\phi_1(x), \ldots, \phi_N(x)$ from the kernel $K(x,t)$. The following three statements clarify this assertion and validate our claim:

1. If $1 \leq j \leq N$, then

$$\begin{aligned}\int_a^b \Delta_{N+1}(x,t)\,\phi_j(t)\,\mathrm{d}t &= \int_a^b K(x,t)\,\phi_j(t)\,\mathrm{d}t \\ &\quad - \sum_{n=1}^{N} \frac{\phi_n(x)}{\lambda_n} \int_a^b \overline{\phi_n(t)}\,\phi_j(t)\,\mathrm{d}t \\ &= \frac{\phi_j(x)}{\lambda_j} - \frac{\phi_j(x)}{\lambda_j} \\ &= 0,\end{aligned}$$

   showing that $\phi_j(x)$ cannot be an eigenfunction of $\Delta_{N+1}(x,t)$.
2. If $\Delta_{N+1}(x,t) \not\equiv 0$, then it has at least one eigenvalue by Proposition 3.2.2. If, for $j \geq N+1$, we subtract the equations

$$\phi_j(x) = \lambda_j \int_a^b K(x,t)\,\phi_j(t)\,\mathrm{d}t$$

and

$$\lambda_j \int_a^b \left(\sum_{n=1}^{N} \frac{\phi_n(x)\overline{\phi_n(t)}}{\lambda_n} \right) \phi_j(t)\,\mathrm{d}t = 0,$$

then we obtain

$$\phi_j(x) = \lambda_j \int_a^b \Delta_{N+1}(x,t)\,\phi_j(t)\,\mathrm{d}t.$$

Thus, each $\phi_j(x)$ is an eigenfunction corresponding to the eigenvalue λ_j of $\Delta_{N+1}(x,t)$.

3. Let ρ be any eigenvalue of the kernel $\Delta_{N+1}(x,t)$ and let $\psi(x) \not\equiv 0$ be the eigenfunction corresponding to it. Then, we have

$$\begin{aligned}\psi(x) &= \rho \int_a^b \Delta_{N+1}(x,t)\,\psi(t)\,\mathrm{d}t \\ &= \rho \int_a^b K(x,t)\,\psi(t)\,\mathrm{d}t - \rho \sum_{n=1}^{N} \frac{\phi_n(x)}{\lambda_n} \int_a^b \overline{\phi_n(t)}\,\psi(t)\,\mathrm{d}t.\end{aligned}$$

If we compute the inner product of $\psi(x)$ with $\phi_j(x)$, with $1 \le j \le N$, then we obtain the relation

$$\langle \psi, \phi_j \rangle = \rho \langle \mathbf{K}\psi, \phi_j \rangle - \frac{\rho}{\lambda_j} \langle \psi, \phi_j \rangle.$$

But since $\mathbf{K}$ is a Hermitian operator, we obtain

$$\langle \mathbf{K}\psi, \phi_j \rangle = \langle \psi, \mathbf{K}^*\phi_j \rangle = \langle \psi, \mathbf{K}\phi_j \rangle = \frac{1}{\lambda_j} \langle \psi, \phi_j \rangle,$$

from which it follows that $\langle \psi, \phi_j \rangle = 0$. By using this fact, we obtain

$$\psi(x) = \rho \int_a^b K(x,t)\,\psi(t)\,\mathrm{d}t,$$

i.e., ρ is an eigenvalue of the kernel $K(x,t)$ and $\psi(x)$ is an eigenfunction corresponding to ρ. As we have shown above, ρ cannot be among the eigenvalues $\lambda_1, \ldots, \lambda_N$. Hence, it must be the case that $\rho = \lambda_{N+k}$ for some $k \ge 1$ and $\psi(x) = \phi_{N+k}$. □

We apply this lemma to obtain the following bilinear expansions:

Theorem 3.3.1 (Representation Formulae). *Let $K(x,t)$ be a complex-valued, continuous, Hermitian kernel defined on the square $Q(a,b)$, with a finite number N of eigenvalues $\lambda_1, \ldots, \lambda_N$, and let $\phi_1(x), \ldots, \phi_N(x)$ be a corresponding system of mutually orthonormal eigenfunctions.*

Then, the following formulae are valid:

1. *The kernel $K(x,t)$ can be represented by the finite bilinear expansion*

$$K(x,t) = \sum_{n=1}^{N} \frac{\phi_n(x)\,\overline{\phi_n(t)}}{\lambda_n}.$$

2. *If $g \in \mathcal{R}^2[a,b]$, then*

$$\mathbf{K}g = \int_a^b K(x,t)\,g(t)\,\mathrm{d}t = \sum_{n=1}^{N} \frac{\langle g,\phi_n\rangle}{\lambda_n}\,\phi_n(x),$$

and

$$\lambda_n \langle \mathbf{K}g,\phi_n\rangle = \langle g,\phi_n\rangle$$

for each $n = 1,\ldots,N$.

3. *For every $m \geq 1$, the iterated kernel $K_m(x,t)$ can be represented by the finite bilinear expansion*

$$K_m(x,t) = \sum_{n=1}^{N} \frac{\phi_n(x)\,\overline{\phi_n(t)}}{\lambda_n^m}.$$

4. *The resolvent kernel $R(x,t;\lambda)$ corresponding to $K(x,t)$ is unique, and it can be represented by the finite bilinear expansion*

$$R(x,t;\lambda) = \sum_{n=1}^{N} \frac{\phi_n(x)\,\overline{\phi_n(t)}}{\lambda_n - \lambda}.$$

Proof. 1. By virtue of Lemma 3.3.1, $K(x,t)$ always has the representation

$$K(x,t) = \sum_{n=1}^{N} \frac{\phi_n(x)\,\overline{\phi_n(t)}}{\lambda_n} + \Delta_{N+1}(x,t).$$

If $\Delta_{N+1}(x,t) \not\equiv 0$, then it must have at least one eigenvalue, say, λ_{N+1}, which must be an eigenvalue of $K(x,t)$. But then $K(x,t)$ would have at least $N+1$ eigenvalues, contrary to hypothesis. Thus, $\Delta_{N+1}(x,t) \equiv 0$.

2. The first equation follows upon direct substitution of the representation for $K(x,t)$ into the integrand. The second equation follows by forming the inner product $\langle \mathbf{K}g,\phi_n\rangle$.

3. The proof follows by induction. One iteration yields

$$\begin{aligned} K_2(x,t) &= \int_a^b \left(\sum_{n=1}^{N} \frac{\phi_n(x)\,\overline{\phi_n(s)}}{\lambda_n}\right)\left(\sum_{j=1}^{N} \frac{\phi_j(s)\,\overline{\phi_j(t)}}{\lambda_j}\right)\mathrm{d}s \\ &= \sum_{n=1}^{N} \frac{\phi_n(x)\,\overline{\phi_n(t)}}{\lambda_n^2}. \end{aligned}$$

Repeated iterations inductively lead to the validity of the representation for all values of m.

4. Since we have arranged the eigenvalues of $K(x,t)$ in order of increasing magnitude, the resolvent series converges if $|\lambda| < |\lambda_1|$. Hence, we can use the representation for $K_m(x,t)$ to obtain

$$\begin{aligned} R(x,t;\lambda) &= \sum_{m=1}^{\infty} \lambda^{m-1} K_m(x,t) \\ &= \sum_{m=1}^{\infty} \lambda^{m-1} \left(\sum_{n=1}^{N} \frac{\phi_n(x)\,\overline{\phi_n(t)}}{\lambda_n^m} \right) \\ &= \sum_{n=1}^{N} \frac{\phi_n(x)\,\overline{\phi_n(t)}}{\lambda_n} \left(\sum_{m=1}^{\infty} (\lambda/\lambda_n)^{m-1} \right) \\ &= \sum_{n=1}^{N} \frac{\phi_n(x)\,\overline{\phi_n(t)}}{\lambda_n} \frac{\lambda_n}{\lambda_n - \lambda} \\ &= \sum_{n=1}^{N} \frac{\phi_n(x)\,\overline{\phi_n(t)}}{\lambda_n - \lambda}. \end{aligned}$$

By the Principle of Analytic Continuation, this representational form is actually valid for all $\lambda \neq \lambda_n$. □

With these representational formulae in hand, it is easy to solve Fredholm integral equations of the first and second kinds with Hermitian separable kernels.

Theorem 3.3.2. *Consider the Fredholm integral equation of the second kind*

$$\phi(x) = f(x) + \lambda \int_a^b K(x,t)\,\phi(t)\,\mathrm{d}t,$$

in which the free term $f(x)$ is continuous on the interval $[a,b]$, and the Hermitian kernel

$$K(x,t) = \sum_{n=1}^{N} \frac{\phi_n(x)\,\overline{\phi_n(t)}}{\lambda_n}$$

is complex-valued and continuous on the square $Q(a,b)$, where $\lambda_1,\ldots,\lambda_N$ are the eigenvalues of the kernel $K(x,t)$ and $\phi_1(x),\ldots,\phi_N(x)$ are the corresponding orthonormal eigenfunctions.

If λ is not an eigenvalue of the kernel, then the unique continuous solution $\phi(x)$ to the integral equation has the representation

$$\phi(x) = f(x) + \lambda \sum_{n=1}^{N} \frac{\langle f, \phi_n \rangle}{\lambda_n - \lambda}\,\phi_n(x).$$

If $\lambda = \lambda_k = \cdots = \lambda_{k+r-1}$ is an eigenvalue of the kernel of multiplicity r with the corresponding eigenfunctions $\phi_k(x), \ldots, \phi_{k+r-1}$, then the integral equation has solutions if and only if $\langle f, \phi_j \rangle = 0$ for every $j = k, \ldots, k+r-1$. In this case, the solution to the integral equation has the representation

$$\phi(x) = f(x) + \lambda_k \sum_j \frac{\langle f, \phi_j \rangle}{\lambda_j - \lambda_k} \phi_j(x) + \sum_{j=k}^{k+r-1} c_j \phi_j(x),$$

where the first sum is taken over the indices j for which $\lambda_j \neq \lambda_k$ and the constants c_j in the second sum are arbitrary.

Proof. If λ is not an eigenvalue of the kernel, then the First Fredholm Theorem for separable kernels as stated in Sect. 1.3 states that the solution to the integral equation is given by

$$\phi(x) = f(x) + \lambda \int_a^b R(x,t;\lambda) f(t)\, \mathrm{d}t.$$

Since $R(x,t;\lambda)$ has the representation given by Theorem 3.3.1(4), the result follows directly.

If λ is an eigenvalue of the kernel, then the Third Fredholm Theorem for separable kernels states that the integral equation has solutions if and only if $f(x)$ is orthogonal to all of the eigenfunctions of the homogeneous adjoint equation corresponding to $\overline{\lambda}$. These are the eigenfunctions of $\overline{K(t,x)}$; however, since $K(x,t)$ is Hermitian, they are the same as the eigenfunctions of $K(x,t)$. Also, by Proposition 3.2.3, λ is real. Thus, for solutions to exist in this case, it is necessary and sufficient that $\langle f, \phi_j \rangle = 0$ for all $j = k, \ldots, k+r-1$.

In Sect. 1.3, it was shown that if these orthogonality conditions are satisfied, then the solution to the inhomogeneous integral equation has the form

$$\phi(x) = f(x) + \lambda_k \phi^{(p)}(x;\lambda_k) + \beta\, \phi^{(h)}(x;\lambda_k),$$

where $f(x) + \phi^{(p)}(x;\lambda_k)$ is a particular solution to the integral equation, β is an arbitrary constant, and $\phi^{(h)}(x;\lambda_k)$ is an arbitrary linear combination of the eigenfunctions corresponding to λ_k. By substituting the representation for the kernel into the integral equation, we obtain

$$\phi(x) = f(x) + \lambda_k \sum_{n=1}^{N} \frac{(\phi, \phi_n)}{\lambda_n} \phi_n(x).$$

If j is an index for which $\lambda_j \neq \lambda_k$, then by forming the scalar product of $\phi(x)$ with $\phi_j(x)$, we obtain

$$\langle \phi, \phi_j \rangle = \langle f, \phi_j \rangle + \lambda_k \frac{\langle \phi, \phi_j \rangle}{\lambda_j}$$

or

$$\langle \phi, \phi_j \rangle = \frac{\lambda_j}{\lambda_j - \lambda_k} \langle f, \phi_j \rangle.$$

By substituting the values into the particular equation, we obtain the first sum in the solution to the equation. The indices j for which $\lambda_j = \lambda_k$ correspond to the second sum in the solution. □

Illustrative Examples

- *Example 1*: Solve the Fredholm integral equation

$$\phi(x) = \mathrm{e}^x + \lambda \int_0^{2\pi} \mathrm{e}^{\mathrm{i}(x-t)}\, \phi(t)\, \mathrm{d}t$$

for all values of λ.

The only eigenvalue of the kernel $K(x,t) = \mathrm{e}^{\mathrm{i}(x-t)}$ is $\lambda_1 = 1/2\pi$, and the corresponding eigenfunction of unit norm is $\phi_1(x) = \mathrm{e}^{\mathrm{i}x}/\sqrt{2\pi}$. A direct computation shows that

$$\langle f, \phi_1 \rangle = \left\langle \mathrm{e}^x, \frac{\mathrm{e}^{\mathrm{i}x}}{\sqrt{2\pi}} \right\rangle = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} \mathrm{e}^{x(1-\mathrm{i})}\, \mathrm{d}x = \frac{1+\mathrm{i}}{2\sqrt{2\pi}} (\mathrm{e}^{2\pi} - 1).$$

If $\lambda \neq 1/2\pi$, then the unique solution to the integral equation is given by

$$\begin{aligned} \phi(x) &= \mathrm{e}^x + \lambda \frac{\langle f, \phi_1 \rangle}{\lambda_1 - \lambda} \phi_1(x) \\ &= \mathrm{e}^x + \lambda \frac{(1+i)(\mathrm{e}^{2\pi} - 1)}{2(1 - 2\pi\lambda)} \mathrm{e}^{\mathrm{i}x}. \end{aligned}$$

If $\lambda = 1/2\pi$, then the equation has no solutions, since $\langle f, \phi_1 \rangle \neq 0$.

- *Example 2*: Solve the Fredholm integral equation

$$\phi(x) = f(x) + \lambda \int_{-2}^{+2} (2x^2 t^2 + xt)\, \phi(t)\, \mathrm{d}t$$

if λ is not an eigenvalue of the kernel.

Upon employing the method presented in Sect. 1.3 for the solution of Fredholm integral equations with separable kernels, we find that the eigenvalues of the kernel are $\lambda_1 = 5/128$ and $\lambda_2 = 3/16$. (Recall that eigenvalues are always enumerated in order of increasing magnitude.) The corresponding eigenfunctions are

$$\phi_1(x) = \frac{\sqrt{5}x^2}{8} \quad \text{and} \quad \phi_2(x) = \frac{\sqrt{3}x}{4},$$

respectively. Observe that $\langle \phi_1, \phi_2 \rangle = 0$, with $\|\phi_1\|_2 = 1$ and $\|\phi_2\|_2 = 1$. The kernel can now be expanded in the structurally revealing form

$$K(x,t) = \frac{\left(\frac{\sqrt{5}x^2}{8}\right)\left(\frac{\sqrt{5}t^2}{8}\right)}{\left(\frac{5}{128}\right)} + \frac{\left(\frac{\sqrt{3}x}{4}\right)\left(\frac{\sqrt{3}t}{4}\right)}{\left(\frac{3}{16}\right)}.$$

Thus, if λ is not an eigenvalue of the kernel, then the solution to the integral equation assumes the form

$$\phi(x) = f(x) + \lambda \frac{\langle f, \phi_1 \rangle}{\frac{5}{128} - \lambda} \frac{\sqrt{5}x^2}{8} + \lambda \frac{\langle f, \phi_2 \rangle}{\frac{3}{16} - \lambda} \frac{\sqrt{3}x}{4},$$

for any free term $f(x)$.

In particular, if $f(x) = \mathrm{e}^x$, then

$$\langle f, \phi_1 \rangle = \left\langle \mathrm{e}^x, \frac{\sqrt{5}x^2}{8} \right\rangle = \frac{\sqrt{5}}{8} \int_{-2}^{+2} x^2 \mathrm{e}^x \, \mathrm{d}x = \frac{\sqrt{5}}{4} (\mathrm{e}^2 - 3\mathrm{e}^{-2}),$$

and

$$\langle f, \phi_2 \rangle = \left\langle \mathrm{e}^x, \frac{\sqrt{3}x}{4} \right\rangle = \frac{\sqrt{3}}{4} \int_{-2}^{+2} x \mathrm{e}^x \, \mathrm{d}x = \frac{\sqrt{3}}{4} (\mathrm{e}^2 + 3\mathrm{e}^{-2}).$$

If we choose $\lambda = 1/32$, then the unique solution to the equation is

$$\phi(x) = \mathrm{e}^x + \frac{5}{8}(\mathrm{e}^2 - 3\mathrm{e}^{-2})x^2 + \frac{3}{80}(\mathrm{e}^2 + 3\mathrm{e}^{-2})x.$$

However, if we choose $\lambda = \lambda_1 = 5/128$ or $\lambda = \lambda_2 = 3/16$, then the integral equation has no solutions whatsoever, since $f(x)$ is not orthogonal to either $\phi_1(x)$ or $\phi_2(x)$.

By way of contrast, the integral equation

$$\phi(x) = x^3 + \frac{5}{128} \int_{-2}^{+2} (2x^2 t^2 + xt) \, \phi(t) \, \mathrm{d}t$$

has an infinite number of solutions of the form

$$\begin{aligned} \phi(x) &= x^3 + \lambda_1 \frac{\langle x^3, \phi_2 \rangle}{\lambda_2 - \lambda_1} \phi_2(x) + c_1 \phi_1 \\ &= x^3 + \frac{12}{19} x + c x^2, \end{aligned}$$

where c is an arbitrary constant, since $\langle x^3, \phi_1 \rangle = 0$.

- *Example 3*: The trace A_2 of the iterated kernel $K_2(x,t)$ was computed in Lemma 3.2.1(2) to be $A_2 = \|K\|_2^2$. But since

$$K_2(x,t) = \sum_1^N \frac{\phi_n(x)\,\overline{\phi(t)}}{\lambda_n^2}$$

by Theorem 3.3.1(3), we also have

$$A_2 = \int_a^b K_2(x,x)\,\mathrm{d}x = \sum_{n=1}^N \frac{1}{\lambda_n^2}.$$

- *Example 4*: The Fredholm integral equation of the *first kind* has the form

$$f(x) = \int_a^b K(x,t)\,\phi(t)\,\mathrm{d}t,$$

where $f(x)$ is continuous on the interval $[a,b]$ and the kernel $K(x,t)$ is continuous on the square $Q(a,b)$. Note that solution $\phi(x)$ does not appear outside of the integrand.

If the kernel $K(x,t)$ is Hermitian and has a finite number of eigenvalues, $\lambda_1,\ldots,\lambda_N$, then it has the representation prescribed in Theorem 3.3.1(1). By substituting it into the above equation, we obtain

$$f(x) = \sum_{n=1}^N \frac{\langle \phi, \phi_n\rangle}{\lambda_n}\,\phi_n(x).$$

Thus, for a solution $\phi(x)$ to the integral equation to exist, it is necessary that $f(x)$ be a linear combination of the eigenfunctions $\phi_1(x),\ldots,\phi_N(x)$ of the kernel $K(x,t)$. Otherwise, the equation has no solution.

If $f(x)$ has the representation

$$f(x) = \sum_{n=1}^N f_n\,\phi_n(x),$$

then a comparison of these representations allows us to conclude that the relations $\langle \phi, \phi_n\rangle = \lambda_n f_n$ hold for all $n = 1,\ldots,N$. Since it is always true that

$$\phi(x) = \sum_{n=1}^N \langle \phi, \phi_n\rangle\,\phi_n(x),$$

we propose that any solution $\phi(x)$ to the integral equation must have the form

$$\phi(x) = \sum_{n=1}^N \lambda_n f_n\,\phi_n(x) + \Phi(x),$$

where $\Phi(x)$ is any function that is orthogonal to all of the eigenfunctions $\phi_n(x)$ on the interval $[a,b]$. Note that

$$\|\phi\|_2^2 = \sum_{n=1}^{N} \lambda_n^2 f_n^2 + \|\Phi\|_2^2,$$

due to the orthonormality of the eigenfunctions.

Some examples illustrate the comments above.

If $K(x,t) = xt + 2x^2t^2$, as in Example 2, then the Fredholm integral equation of the first kind

$$\cos x = \int_{-2}^{+2} K(x,t)\,\phi(t)\,\mathrm{d}t = \left(\int_{-2}^{+2} t\phi(t)\,\mathrm{d}t\right) x + \left(\int_{-2}^{+2} 2t^2\phi(t)\,\mathrm{d}t\right) x^2$$

has no solutions, since $\cos x$ is not a linear combination of x and x^2.

However, the integral equation

$$4x^2 + 3x = \int_{-2}^{+2} K(x,t)\,\phi(t)\,\mathrm{d}t$$

has solutions of the form

$$\phi(x) = \lambda_1 f_1\,\phi_1(x) + \lambda_2 f_2\,\phi_2(x) + \Phi(x),$$

since $4x^2 + 3x$ can be expressed in terms of the eigenvalues $\lambda_1 = 5/128$ and $\lambda_2 = 3/16$, and the eigenfunctions $\phi_1(x) = \sqrt{5}x^2/8$ and $\phi_2(x) = \sqrt{3}x/4$. Indeed, since

$$4x^2 + 3x = \frac{32}{\sqrt{5}}\,\phi_1(x) + \frac{12}{\sqrt{3}}\,\phi_2(x),$$

we have $f_1 = 32/\sqrt{5}$ and $f_2 = 12/\sqrt{3}$. By substituting these values into the general form of the solution to the integral equation, we obtain

$$\phi(x) = \frac{5}{32}x^2 + \frac{9}{16}x + \Phi(x),$$

where $\Phi(x)$ is any continuous function which is orthogonal to both x and x^2 on the interval $[-2,+2]$.

Section 3.3 Exercises

1. (*Thought Question*) Are the eigenvalues and eigenfunctions of a kernel dependent on the domain of its definition? Let $K(x,t)$ be a complex-valued, nonvanishing, continuous, Hermitian kernel, and let $0 < a < b < c$. If $K_{ab}(x,t) = K(x,t)$ on $Q(a,b)$ and $K_{ac}(x,t) = K(x,t)$ on $Q(a,c)$, are the

eigenvalues and eigenfunctions of $K_{ab}(x,t)$ identical to those of $K_{ac}(x,t)$? If two eigenfunctions of $K_{ab}(x,t)$ are orthogonal on the interval $[a,b]$, are they orthogonal on the interval $[a,c]$?

2. Recall that the eigenvalues are always arranged in order of increasing magnitude. Use Example 3 to show that

$$\lambda_1^2 \leq \frac{\|K\|_2^2}{N} \leq \lambda_N^2,$$

thereby furnishing an upper bound on $|\lambda_1|$ and a lower bound on $|\lambda_N|$.

3. Let $\{L_n(x)\}_{n=0}^{\infty}$ denote the mutually orthonormal system of Legendre polynomials described in Sect. 3.1.

   (a) Let $p(x)$ be an arbitrary polynomial of degree N. Explain why there must exist constants p_n such that

$$p(x) = \sum_{n=0}^{N} p_n L_n(x),$$

   and derive a formula for each p_n.

   (b) Solve the integral equation

$$\phi(x) = q(x) + \lambda \int_{-1}^{+1} K(x,t)\,\phi(t)\,\mathrm{d}t,$$

   where $q(x) = x^4 + x^2 + 1$,

$$K(x,t) = \sum_{n=0}^{4} \frac{L_n(x)\,L_n(t)}{n+1},$$

   and λ is not an eigenvalue of $K(x,t)$.

   Hint: Begin by expressing $q(x)$ as a linear combination of Legendre polynomials, as done in part (a).

   (c) Solve the Fredholm integral equation of the first kind

$$x^4 + x^2 + 1 = \int_{-1}^{+1} K(x,t)\,\psi(t)\,\mathrm{d}t,$$

   where $K(x,t)$ is the kernel given in part (b).

3.4 The Hilbert–Schmidt Theorem

The finite bilinear expansions given in Theorem 3.3.1 allowed for the solution of Fredholm integral equations of the second kind with Hermitian separable kernels. It is a natural question to ask whether these expansions remain valid if the upper

index of summation N is replaced with ∞. In general, the response is negative, since the series

$$\sum_{n=1}^{\infty} \frac{\phi_n(x)\,\overline{\phi_n(t)}}{\lambda_n}$$

need not converge. However, all is not lost. The crucial theorem for us in this regard is the Hilbert–Schmidt Theorem stated below.

Although we assume in the Hilbert–Schmidt Theorem that a Fourier series of the form

$$g(x) \sim \sum_n \langle g, \phi_n \rangle \, \phi_n(x) = \sum_n g_n \phi_n(x)$$

can be constructed, we do not assume any type of convergence for it. Neither do we assume that the set of orthonormal eigenfunctions $\{\phi_n(x)\}$ is complete.

Theorem 3.4.1 (The Hilbert–Schmidt Theorem). *Let $K(x,t)$ be a complex-valued, continuous, Hermitian kernel defined on the square $Q(a,b)$. Suppose that the kernel $K(x,t)$ has an infinite number of eigenvalues $\lambda_1, \lambda_2, \ldots$ and that the corresponding orthonormal eigenfunctions are $\phi_1(x), \phi_2(x), \ldots$*

For $g \in \mathcal{R}^2[a,b]$, define

$$f(x) = \mathbf{K}g = \int_a^b K(x,t)\, g(t)\, \mathrm{d}t.$$

Then, $f(x)$ can be expanded in the absolutely and uniformly convergent Fourier series

$$f(x) = \sum_{n=1}^{\infty} f_n \, \phi_n(x) = \sum_{n=1}^{\infty} \frac{g_n}{\lambda_n} \, \phi_n(x),$$

where $f_n = \langle f, \phi_n \rangle$ and $g_n = \langle g, \phi_n \rangle$ are the Fourier coefficients of $f(x)$ and $g(x)$, respectively. Also, $f \in \mathcal{R}^2[a,b]$.

Proof. Since $K(x,t)$ is Hermitian, each λ_n is real by Proposition 3.2.3. The Fourier coefficients of $f(x)$ are related to those of $g(x)$, since

$$f_n = \langle f, \phi_n \rangle = \langle \mathbf{K}g, \phi_n \rangle = \langle g, \mathbf{K}\phi_n \rangle = \left\langle g, \frac{\phi_n}{\lambda_n} \right\rangle = \frac{1}{\lambda_n} \langle g, \phi_n \rangle = \frac{g_n}{\lambda_n}.$$

In order to prove the theorem, we will first show that the Fourier series converges absolutely and uniformly to some function $\sigma(x)$, and then show that $f(x) = \sigma(x)$.

To establish absolute and uniform convergence of the Fourier series, we show that the sequence

$$\sigma_N(x) = \sum_{n=1}^{N} \frac{g_n}{\lambda_n} \, \phi_n(x)$$

of its partial sums satisfies the Cauchy criterion. As a consequence of the Cauchy–Schwarz inequality, we have

$$\begin{aligned}\left|\sigma_{N+p}(x)-\sigma_N(x)\right|^2 &= \left|\sum_{n=N+1}^{N+p} g_n\,\frac{\phi_n(x)}{\lambda_n}\right|^2 \\ &\le \left(\sum_{n=N+1}^{N+p} g_n^2\right)\left(\sum_{n=N+1}^{N+p}\frac{|\phi_n(x)|^2}{\lambda_n^2}\right) \\ &\le \left(\sum_{n=N+1}^{N+p} g_n^2\right)\left(\sum_{n=1}^{\infty}\frac{|\phi_n(x)|^2}{\lambda_n^2}\right)\end{aligned}$$

for any positive integer p. For N large enough, the first factor on the right can be made arbitrarily small, since

$$\sum_{n=1}^{\infty} g_n^2 = \sum_{n=1}^{\infty} |\langle g,\phi_n\rangle|^2 \le \|g\|_2^2 < +\infty$$

by Bessel's inequality and the assumption that $g \in \mathcal{R}^2[a,b]$. The second factor on the right is bounded, again as a consequence of Bessel's inequality. For each fixed $x \in [a,b]$, consider the continuous function $k_x(s) = \overline{K(x,s)}$. The Fourier coefficients of $k_x(s)$ with respect to the orthonormal system $\{\overline{\phi_n(s)}\}$ are given by

$$\langle k_x, \overline{\phi_n}\rangle = \int_a^b K(x,s)\,\phi_n(s)\,\mathrm{d}s = \frac{\phi_n(x)}{\lambda_n}.$$

Hence,

$$\sum_{n=1}^{\infty}\frac{|\phi_n(x)|^2}{\lambda_n^2} = \sum_{n=1}^{\infty}\left|\langle k_x,\overline{\phi_n}\rangle\right|^2 \le \|k_x\|_2^2 < +\infty. \tag{3.2}$$

Thus, the sequence $\{\sigma_n(x)\}_{N=1}^{\infty}$ satisfies the Cauchy criterion, implying that there exists a continuous function $\sigma \in \mathcal{R}^2[a,b]$ such that

$$\sigma(x) = \lim_{N\to\infty}\sigma_n(x).$$

Now we must show that $f(x) = \sigma(x)$. To do this, we note that

$$\|f-\sigma\|_2 \le \|f-\sigma_N\|_2 + \|\sigma_N-\sigma\|_2$$

as an application of the triangle inequality and that the left-hand side of this inequality is independent of N. To conclude the proof, it must be shown that $\|\sigma_N-\sigma\|_2 \to 0$ and that $\|f-\sigma_n\|_2 \to 0$ as $N \to +\infty$.

Given an arbitrary $\varepsilon > 0$, there exists an integer N_ε such that

$$|\sigma_N(x) - \sigma(x)| < \frac{\varepsilon}{2\sqrt{b-a}}$$

for all $N \geq N_\varepsilon$, since the Fourier series converges uniformly. It follows from this inequality that

$$\|\sigma_N - \sigma\|_2^2 = \int_a^b |\sigma_N(x) - \sigma(x)|^2 \, \mathrm{d}x < \frac{\varepsilon^2}{4},$$

or that $\|\sigma_N - \sigma\|_2 < \varepsilon/2$ whenever $N \geq N_\varepsilon$.

We turn to $\|f - \sigma_N\|_2$. Observe first that

$$\begin{aligned} f(x) - \sigma_N(x) &= \mathbf{K}g - \sum_{n=1}^{N} \frac{g_n}{\lambda_n} \phi_n(x) \\ &= \int_a^b K(x,t)\, g(t)\, \mathrm{d}t - \sum_{n=1}^{N} \frac{\langle g, \phi_n \rangle}{\lambda_n} \phi_n(x) \\ &= \int_a^b K(x,t)\, g(t)\, \mathrm{d}t - \sum_{n=1}^{N} \frac{1}{\lambda_n} \left(\int_a^b g(t)\, \overline{\phi_n(t)}\, \mathrm{d}t \right) \phi_n(x) \\ &= \int_a^b \left(K(x,t) - \sum_{n=1}^{N} \frac{\phi_n(x)\overline{\phi_n(t)}}{\lambda_n} \right) g(t)\, \mathrm{d}t \\ &= \int_a^b \Delta_{N+1}(x,t)\, g(t)\, \mathrm{d}t. \end{aligned}$$

We recognize $\Delta_{N+1}(x,t)$ here as the truncated, complex-valued, continuous, and Hermitian kernel defined in the statement of Lemma 3.3.1. If we define the Hermitian operator

$$\mathbf{D}_{N+1} : g \in \mathcal{R}^2[a,b] \to \mathbf{D}_{N+1}\, g = \int_a^b \Delta_{N+1}(x,t)\, g(t)\, \mathrm{d}t,$$

then

$$\|f - \sigma_N\|_2^2 = \|\mathbf{D}_{N+1}\, g\|_2^2 = \langle \mathbf{D}_{N+1}\, g, \mathbf{D}_{N+1}\, g \rangle = \langle g, \mathbf{D}_{N+1}^2\, g \rangle.$$

The operator $\mathbf{D}_{N+1}^2$ is also Hermitian, and it corresponds to the iteration of the kernel $\Delta_{N+1}(x,t)$ with itself. By Lemma 3.3.1, its least eigenvalue is λ_{N+1}^2, and by Proposition 3.2.1, we have

$$\frac{1}{\lambda_{N+1}^2} = \max_{\phi} \frac{\langle \phi, \mathbf{D}_{N+1}^2\, \phi \rangle}{\langle \phi, \phi \rangle}.$$

For any $g \in \mathcal{R}^2[a,b]$, it follows that

$$\frac{1}{\lambda_{N+1}^2} \geq \frac{\langle g, \mathbf{D}_{N+1}^2\, g \rangle}{\langle g, g \rangle} = \frac{\|f - \sigma_N\|_2^2}{\langle g, g \rangle}.$$

By the Fourth Fredholm Theorem, $\lambda_N \to +\infty$ as $n \to +\infty$. Therefore, there exists an M_ε such that

$$\frac{\langle g,g\rangle}{\lambda_{N+1}^2} < \frac{\varepsilon^2}{4}$$

for all $N \geq M_\varepsilon$. By combining these two inequalities, we obtain $\|f-\sigma_N\|_2 < \varepsilon/2$ for all $N \geq M_\varepsilon$. The proof is now complete, since

$$\|f-\sigma\|_2 \leq \|f-\sigma_N\|_2 + \|\sigma_N-\sigma\|_2 < \frac{\varepsilon}{2}+\frac{\varepsilon}{2} = \varepsilon$$

for all $N \geq \max\{N_\varepsilon, M_\varepsilon\}$. But since ε is arbitrary, it must now be the case that $\|f-\sigma\|_2 = 0$.

Finally, $f \in \mathcal{R}^2[a,b]$, since $\|f\|_2 = \|\mathbf{K}g\|_2 \leq \|\mathbf{K}\| \cdot \|g\|_2$. □

Corollary 3.4.1 (Hilbert's Formula). *Let $K(x,t)$ have the same properties as stated in Theorem 3.4.1.*

If $g,h \in \mathcal{R}^2[a,b]$, then

$$\langle \mathbf{K}g,h\rangle = \sum_{n=1}^{\infty} \frac{1}{\lambda_n}\langle g,\phi_n\rangle\langle \phi_n,h\rangle.$$

In particular, if $g=h$, then

$$\langle \mathbf{K}g,g\rangle = \sum_{n=1}^{\infty} \frac{1}{\lambda_n}|\langle g,\phi_n\rangle|^2.$$

Proof. Since the Fourier series converges absolutely, it can be integrated term by term. □

Corollary 3.4.2. *Let $K(x,t)$ have the same properties as stated in Theorem 3.4.1. Then, for $m \geq 2$, the iterated kernel $K_m(x,t)$ has the bilinear expansion*

$$K_m(x,t) = \sum_{n=1}^{\infty} \frac{\phi_n(x)\overline{\phi_n(t)}}{\lambda_n^m}.$$

Consequently, the trace A_m of $K_m(x,t)$ is given by

$$A_m = \int_a^b K_m(x,x)\,\mathrm{d}x = \sum_{n=1}^{\infty} \frac{1}{\lambda_n^m}.$$

Proof. The proof proceeds by induction. For fixed $s \in [a,b]$, let $g(t) = K(t,s)$. Then, since λ_n is real and $K(x,t)$ is Hermitian, the Fourier coefficients of $g(t)$ are computed to be

$$g_n = \langle g,\phi_n\rangle = \int_a^b K(t,s)\overline{\phi_n(t)}\,\mathrm{d}t = \int_a^b \overline{K(s,t)}\,\overline{\phi_n(t)}\,\mathrm{d}t = \frac{\overline{\phi_n(s)}}{\lambda_n}.$$

On the one hand, we have

$$\mathbf{K}g = \sum_{n=1}^{\infty} \frac{g_n}{\lambda_n}\,\phi_n(x) = \sum_{n=1}^{\infty} \frac{1}{\lambda_n^2}\,\phi_n(x)\overline{\phi_n(s)},$$

while on the other hand, we have

$$\mathbf{K}g = \int_a^b K(x,t)\,K(t,s)\,\mathrm{d}t = K_2(x,s).$$

Hence, the bilinear expansion is valid if $m = 2$. In this case, it follows from the theorem that the convergence is absolute and uniform in the variable x for each fixed s and vice versa.

Now suppose that the expansion is valid if $m = 2,\ldots,M$. By setting $g(t) = K_M(t,s)$ in the above equations, similar calculations show that the required expansion is also valid if $m = M+1$.

Actually, for $m \geq 3$, the convergence is uniform with respect to both of the variables x and s. Let p be a positive integer and recall that the eigenvalues are arranged in order of increasing magnitude. Consider the inequalities

$$\begin{aligned}
\left|\sum_{n=N+1}^{N+p} \frac{\phi_n(x)\,\overline{\phi_n(s)}}{\lambda_n^m}\right|^2 &\leq \frac{1}{\left|\lambda_{N+1}^{m-2}\right|^2}\left(\sum_{n=N+1}^{N+p}\left|\frac{\phi_n(x)}{\lambda_n}\right|\left|\frac{\phi_n(s)}{\lambda_n}\right|\right)^2 \\
&\leq \frac{1}{\left|\lambda_{N+1}^{m-2}\right|^2}\left(\sum_{n=1}^{\infty}\left|\frac{\phi_n(x)}{\lambda_n}\right|\left|\frac{\phi_n(s)}{\lambda_n}\right|\right)^2 \\
&\leq \frac{1}{\left|\lambda_{N+1}^{m-2}\right|^2}\left(\sum_{n=1}^{\infty}\frac{|\phi_n(x)|^2}{\lambda_n^2}\right)\left(\sum_{n=1}^{\infty}\frac{|\phi_n(s)|^2}{\lambda_n^2}\right) \\
&\leq \frac{1}{\left|\lambda_{N+1}^{m-2}\right|^2}\,\|k_x\|_2^2\,\|k_s\|_2^2.
\end{aligned}$$

Since the norms of k_x and k_s are finite by Eq. (3.2) and $|\lambda_n| \to 0$ as $n \to +\infty$, it follows that the bilinear series satisfies the Cauchy criterion and therefore converges absolutely and uniformly on the square $Q(a,b)$. □

Corollary 3.4.3 (Resolvent Kernel). *Let $K(x,t)$ have the same properties as stated in Theorem 3.4.1. Then the resolvent kernel $R(x,t;\lambda)$ corresponding to $K(x,t)$ has the infinite bilinear expansion*

$$R(x,t;\lambda) = K(x,t) + \lambda \sum_{n=1}^{\infty} \frac{1}{\lambda_n\,(\lambda_n - \lambda)}\,\phi_n(x)\,\overline{\phi_n(t)}$$

which converges absolutely and uniformly if $\lambda \neq \lambda_n$ for all $n \geq 1$.

Proof. For $x,t \in Q(a,b)$, the resolvent series converges absolutely whenever $|\lambda| < |\lambda_1|$ by the Theorem of Successive Approximation stated in Sect. 2.3.

When constructing the bilinear series for the resolvent, one must consider the fact that the bilinear series for the first term $K_1(x,t) = K(x,t)$ need not converge. As a consequence of the previous corollary, we have

$$\begin{aligned}
R(x,t;\lambda) &= \sum_{m=1}^{\infty} \lambda^{m-1} K_m(x,t) \\
&= K(x,t) + \sum_{m=2}^{\infty} \lambda^{m-1} K_m(x,t) \\
&= K(x,t) + \frac{1}{\lambda} \sum_{m=2}^{\infty} \sum_{n=1}^{\infty} \left(\frac{\lambda}{\lambda_n}\right)^m \phi_n(x)\,\overline{\phi_n(t)} \\
&= K(x,t) + \frac{1}{\lambda} \sum_{n=1}^{\infty} \left[\sum_{m=2}^{\infty} \left(\frac{\lambda}{\lambda_n}\right)^m\right] \phi_n(x)\,\overline{\phi_n(t)} \\
&= K(x,t) + \frac{1}{\lambda} \sum_{n=1}^{\infty} \left[\frac{(\lambda/\lambda_n)^2}{1-(\lambda/\lambda_n)}\right] \phi_n(x)\,\overline{\phi_n(t)} \\
&= K(x,t) + \lambda \sum_{n=1}^{\infty} \frac{1}{\lambda_n\,(\lambda_n - \lambda)}\, \phi_n(x)\,\overline{\phi_n(t)}.
\end{aligned}$$

By the Principle of Analytic Continuation, this representational form is actually valid for all $\lambda \neq \lambda_n$ for any $n \geq 1$. □

Note that the form of the resolvent in this corollary differs from the form of the finite bilinear expansion for the resolvent as stated in Theorem 3.3.1(4). This apparent difference is due to the fact that the series

$$K(x,t) = \sum_{n=1}^{\infty} \frac{\phi_n(x)\,\overline{\phi_n(t)}}{\lambda_n}$$

need not converge absolutely and uniformly on $Q(a,b)$. If the convergence were absolute and uniform, then the inclusion of this representational form results in the formula

$$R(x,t;\lambda) = \sum_{n=1}^{\infty} \frac{\phi_n(x)\,\overline{\phi_n(t)}}{\lambda_n - \lambda}$$

which is the expected analogous representation.

Given this observation, it is of interest to discover a type of kernel with an infinite bilinear expansion that does converge absolutely and uniformly. Mercer discovered one such type of kernel.

Let $K(x,t)$ be a nonvanishing, complex-valued, continuous Hermitian kernel. Then $K(x,t)$ is called *positive* if $\langle \mathbf{K}\phi,\phi\rangle \geq 0$ for all $\phi \in \mathcal{R}[a,b]$. If, in addition, $\langle \mathbf{K}\phi,\phi\rangle > 0$ if and only if $\langle \phi,\phi\rangle > 0$, then $K(x,t)$ is called *positive definite*.[1] If the orthonormal system $\{\phi_n(x)\}$ is not complete, there may exist a nonvanishing function $\phi(x)$ for which $\langle \mathbf{K}\phi,\phi\rangle = 0$.

Positive kernels have some noteworthy properties.

Proposition 3.4.1. *Let $K(x,t)$ be a nonvanishing, complex-valued, continuous Hermitian kernel defined on the square $Q(a,b)$.*

If $K(x,t)$ is positive, then the following statements are true:

1. *The eigenvalues λ_n of $K(x,t)$ are all positive.*
2. *For every $x \in [a,b]$, the inequality $K(x,x) \geq 0$ holds.*
3. *The Fredholm operator defined by*

$$\mathbf{D}_{N+1}\, g(x) = \int_a^b \Delta_{N+1}(x,t)\, g(t)\, \mathrm{d}t,$$

where $\Delta_{N+1}(x,t)$ is the truncated kernel

$$\Delta_{N+1}(x,t) = K(x,t) - \sum_{n=1}^{N} \frac{\phi_n(x)\,\overline{\phi_n(t)}}{\lambda_n},$$

satisfies the properties

$$\mathbf{D}_{N+1}\,\phi(x) = \sum_{N+1}^{\infty} \frac{1}{\lambda_n}\,\langle \phi,\phi_n\rangle\,\phi_n(x)$$

and

$$\langle \mathbf{D}_{N+1}\phi,\phi\rangle = \sum_{N+1}^{\infty} \frac{1}{\lambda_n}\,|\langle \phi,\phi_n\rangle|^2$$

Consequently, $\Delta_{N+1}(x,t)$ is a positive kernel.

4. *For every $N \geq 1$ and $x \in [a,b]$, we have*

$$\sum_{n=1}^{N} \frac{1}{\lambda_n}\,|\phi_n(x)|^2 \leq K(x,x).$$

Proof. 1. If $\lambda_n \mathbf{K}\phi_n = \phi_n$, then $\lambda_n \langle \mathbf{K}\phi_n,\phi_n\rangle = \langle \phi_n,\phi_n\rangle > 0$. Thus, $\lambda_n > 0$.

2. The proof proceeds by contradiction. Let $c \in (a,b)$, and assume that $K(c,c) < 0$. Since $K(x,t)$ is continuous on $Q(a,b)$, there exists a (square) neighborhood

$$S(c,\delta) = \{(x,t)\colon c-\delta \leq x \leq c+\delta \text{ and } c-\delta \leq t \leq c+\delta\}$$

[1] This terminology varies slightly in the literature.

of the point $(c,c) \in Q(a,b)$ such that $\mathrm{Re}\{K(x,t)\} < 0$ on $S(c,\delta)$, provided that δ is sufficiently small. Define the unit pulse function

$$p(x;c,\delta) = \begin{cases} 1, & \text{if } c-\delta \le x \le c+\delta \\ 0, & \text{if } a \le x < c-\delta \text{ and } c+\delta < x \le b. \end{cases}$$

Since $\langle \mathbf{K}p, p\rangle$ is real-valued by Proposition 3.2.1(5), we have

$$\begin{aligned} \langle \mathbf{K}p, p\rangle &= \int_a^b \int_a^b K(x,t)\, p(t)\, \overline{p(x)}\, \mathrm{d}x\, \mathrm{d}t \\ &= \iint_{S(c,\delta)} \mathrm{Re}\{K(x,t)\}\, \mathrm{d}x\, \mathrm{d}t < 0, \end{aligned}$$

contrary to the assumption that $K(x,t)$ is a positive kernel. We also have $K(a,a) \ge 0$ and $K(b,b) \ge 0$ by a continuity argument.

3. By Lemma 3.3.1, $\Delta_{N+1}(x,t)$ is continuous on $Q(a,b)$, and its eigenvalues are the eigenvalues of $K(x,t)$ for which $n \ge N+1$. If $\phi \in \mathcal{R}^2[a,b]$, then the representation for $\mathbf{D}_{N+1}\phi$ follows from the Hilbert–Schmidt Theorem. The representation for $\langle \mathbf{D}_{N+1}\phi, \phi\rangle$ is then an application of Corollary 3.4.1 (Hilbert's formula). Since the eigenvalues λ_n are positive for $n \ge N+1$, we conclude that $\Delta_{N+1}(x,t)$ is a positive kernel.
4. Since $\Delta_{N+1}(x,t)$ is a positive kernel, we have

$$\Delta_{N+1}(x,x) = K(x,x) - \sum_{n=1}^{N} \frac{1}{\lambda_n} |\phi_n(x)|^2 \ge 0.$$

The result follows directly from this observation. □

We have noted above that the infinite bilinear expansion of a Hermitian kernel need not converge. However, it is still reasonable to expect that there would exist conditions under which the infinite bilinear expansion of a Hermitian kernel $K(x,t)$ would converge, since the Fourier development has the expected form

$$K(x,t) \sim \sum_{n=1}^{\infty} \frac{\phi_n(x)\, \overline{\phi_n(t)}}{\lambda_n}.$$

To see this, suppose that

$$K(x,t) \sim \sum_{n=1}^{\infty} \alpha_n\, \phi_n(x)$$

for a fixed value of t. Then,

$$\alpha_n = \int_a^b K(x,t)\,\overline{\phi_n(x)}\,\mathrm{d}x$$

$$= \int_a^b \overline{K(t,x)}\,\overline{\phi_n(x)}\,\mathrm{d}x$$

$$= \frac{\overline{\phi_n(t)}}{\lambda_n}.$$

We are now in a position to establish the following result.

Theorem 3.4.2 (Mercer's Theorem). *Let $K(x,t)$ be a nonvanishing, complex-valued, continuous, Hermitian kernel.*

If $K(x,t)$ is positive, then it can be expanded in the infinite bilinear series

$$K(x,t) = \sum_{n=1}^{\infty} \frac{\phi_n(x)\,\overline{\phi_n(t)}}{\lambda_n}$$

which converges absolutely and uniformly on $Q(a,b)$.

Proof. By Proposition 3.4.1(4), the infinite series

$$\sum_{n=1}^{\infty} \frac{1}{\lambda_n} |\phi_n(x)|^2 \tag{3.3}$$

converges for each $x \in [a,b]$, and its sum is no greater than $K(x,x)$.

Note: It is possible that $K(x_0,x_0) = 0$ for some value of x_0. For example, if $\phi_n(x) = \sin(nx)$, then $\phi_n(\pi) = 0$, which directly implies that $K(\pi,\pi) = 0$. However, it is not possible for $K(x,x) \equiv 0$ on the interval $[a,b]$. If that were the case, then we would also have $\phi_n(x) \equiv 0$ on $[a,b]$ for all n.

Now let $\varepsilon > 0$ be arbitrary and define

$$B = \max_{a \le x \le b} K(x,x).$$

Since the series (3.3) converges for each $x \in [a,b]$, there exists an integer $N_x(\varepsilon)$ depending on x and ε such that

$$\sum_{n=m}^{p} \frac{1}{\lambda_n} |\phi_n(x)|^2 < \frac{\varepsilon}{B} \tag{3.4}$$

for all $p \ge m \ge N_x(\varepsilon)$.

We now turn to the bilinear series. As an application of the Cauchy–Schwarz inequality, we have

$$\begin{aligned}\left|\sum_{n=m}^{p}\phi_n(x)\,\overline{\phi_n(t)}\right|^2 &\le \left(\sum_{n=m}^{p}\left|\phi_n(x)\,\overline{\phi_n(t)}\right|\right)^2\\ &\le \left(\sum_{n=m}^{p}\frac{1}{\lambda_n}\,|\phi_n(x)|^2\right)\left(\sum_{n=m}^{p}\frac{1}{\lambda_n}\,|\phi_n(t)|^2\right)\\ &\le \left(\sum_{n=m}^{p}\frac{1}{\lambda_n}\,|\phi_n(x)|^2\right)K(t,t)\\ &< \frac{\varepsilon}{B}\cdot B\\ &= \varepsilon \end{aligned} \tag{3.5}$$

for all $p \ge m \ge N_x(\varepsilon)$. Thus, for each fixed $x \in [a,b]$, the infinite bilinear series converges uniformly to a function $L(x,t)$ that is continuous with respect to the variable t.

Next, we wish to show that $L(x,t) = K(x,t)$ on the square $Q(a,b)$. Let $\phi \in \mathcal{R}^2[a,b]$. Then, as a consequence of our previous remarks and the Hilbert–Schmidt Theorem, we have

$$\begin{aligned}\int_a^b (L(x,t)-K(x,t))\,\phi(t)\,\mathrm{d}t &= \int_a^b L(x,t)\,\phi(t)\,\mathrm{d}t - \int_a^b K(x,t)\,\phi(t)\,\mathrm{d}t\\ &= \int_a^b \left(\sum_{n=1}^{\infty}\frac{1}{\lambda_n}\,\phi_n(x)\,\overline{\phi_n(t)}\right)\phi(t)\,\mathrm{d}t - \sum_{n=1}^{\infty}\frac{1}{\lambda_n}\langle\phi,\phi_n\rangle\,\phi_n(x)\\ &= \sum_{n=1}^{\infty}\frac{1}{\lambda_n}\,\phi_n(x)\left(\int_a^b \phi(t)\,\overline{\phi_n(x)}\,\mathrm{d}t\right) - \sum_{n=1}^{\infty}\frac{1}{\lambda_n}\langle\phi,\phi_n\rangle\,\phi_n(x)\\ &= 0.\end{aligned}$$

In particular, if we choose the continuous function $\phi(t) = \overline{L(x,t)-K(x,t)}$, then we obtain

$$\|L(x,t)-K(x,t)\|_2^2 = \int_a^b |L(x,t)-K(x,t)|^2\,\mathrm{d}t = 0.$$

Since both $L(x,t)$ and $K(x,t)$ are continuous functions of t on the interval $[a,b]$, it now follows that $L(x,t) = K(x,t)$ for each fixed $x \in [a,b]$. Hence,

$$L(x,t) = K(x,t) = \sum_{n=1}^{\infty}\frac{\phi_n(x)\,\overline{\phi_n(t)}}{\lambda_n}$$

on the square $Q(a,b)$.

It remains to show that the convergence of the bilinear series is uniform in the variable x. By choosing $t = x$, we obtain

$$K(x,x) = \sum_{n=1}^{\infty}\frac{1}{\lambda_n}\,|\phi_n(x)|^2. \tag{3.6}$$

Since the limit function $K(x,x)$ is continuous on the interval $[a,b]$, this series converges *uniformly* there by Dini's theorem.[2] This means that there exists an integer $N(\varepsilon)$ such that the inequality (3.4) holds for all $p \geq m \geq N(\varepsilon)$ *independently of* x. As a direct consequence, the inequality (3.5) also holds under these conditions. Thus, the infinite bilinear series converges absolutely and uniformly to $K(x,t)$, and the proof is complete. □

Corollary 3.4.4. *Let $K(x,t)$ have the same properties as stated in Theorem 3.4.2. Then the trace A_1 of the kernel $K(x,t)$ is given by*

$$A_1 = \int_a^b K(x,x)\,\mathrm{d}x = \sum_{n=1}^{\infty} \frac{1}{\lambda_n}.$$

Theorem 3.4.3. *Consider the Fredholm integral equation of the second kind*

$$\phi(x) = f(x) + \lambda \int_a^b K(x,t)\,\phi(t)\,\mathrm{d}t,$$

where the free term $f(x)$ is continuous on the interval $[a,b]$, and the kernel $K(x,t)$ is nonvanishing, complex-valued, continuous, and Hermitian on the square $Q(a,b)$. Let $\lambda_1, \lambda_2, \ldots$ be the eigenvalues of the kernel arranged in order of increasing magnitude, and let $\phi_1(x), \phi_2(x), \ldots$ be the mutually orthonormal system of eigenfunctions corresponding to them.

If λ is not an eigenvalue of the kernel, then the unique continuous solution $\phi(x)$ to the integral equation has the representation

$$\phi(x) = f(x) + \lambda \sum_{n=1}^{\infty} \frac{\langle f, \phi_n \rangle}{\lambda_n - \lambda}\,\phi_n(x).$$

If $\lambda = \lambda_k = \cdots = \lambda_{k+r-1}$ is an eigenvalue of the kernel of multiplicity r with the corresponding eigenfunctions $\phi_k(x), \ldots, \phi_{k+r-1}$, then the integral equation has solutions if and only if $\langle f, \phi_j \rangle = 0$ for every $j = k, \ldots, k+r-1$. In this case, the solution to the integral equation has the representation

$$\phi(x) = f(x) + \lambda_k \sum_{j} \frac{\langle f, \phi_j \rangle}{\lambda_j - \lambda_k}\,\phi_j(x) + \sum_{j=k}^{k+r-1} c_j\,\phi_j(x),$$

where the first sum is taken over the indices j for which $\lambda_j \neq \lambda_k$ and the constants c_j in the second sum are arbitrary.

[2] Dini's theorem states that an increasing sequence of functions that converges pointwise to a continuous function on a closed interval actually converges uniformly there.

Proof. Under the given hypotheses, the integral equation can be reformulated as

$$\phi(x) = f(x) + \lambda \sum_{n=1}^{\infty} \frac{\langle \phi, \phi_n \rangle}{\lambda_n} \phi_n(x)$$

as a consequence of the Hilbert–Schmidt Theorem. Recall that the convergence here is both absolute and uniform. By taking the inner product of each of the terms in this equation with $\phi_n(x)$, we directly obtain

$$\langle \phi, \phi_n \rangle = \langle f, \phi_n \rangle + \lambda \frac{\langle \phi, \phi_n \rangle}{\lambda_n}$$

or, equivalently,

$$\langle \phi, \phi_n \rangle = \frac{\langle f, \phi_n \rangle \lambda_n}{\lambda_n - \lambda}.$$

The result follows upon substitution.

The solution can be written in an alternate form by replacing $\langle f, \phi_n \rangle$ with its integral definition. Since

$$\begin{aligned}
\sum_{n=1}^{\infty} \frac{\langle f, \phi_n \rangle}{\lambda_n - \lambda} \phi_n(x) &= \sum_{n=1}^{\infty} \left(\int_a^b f(t) \overline{\phi_n(t)} \, \mathrm{d}t \right) \frac{\phi_n(x)}{\lambda_n - \lambda} \\
&= \int_a^b \left(\sum_{n=1}^{\infty} \frac{\phi_n(x) \overline{\phi_n(t)}}{\lambda_n - \lambda} \right) f(t) \, \mathrm{d}t \\
&= \int_a^b R(x,t;\lambda) f(t) \, \mathrm{d}t,
\end{aligned}$$

the solution can be written in terms of the resolvent, as was done in the First Fredholm Theorem, which states that the solution to the integral equation is given by

$$\phi(x) = f(x) + \lambda \int_a^b R(x,t;\lambda) f(t) \, \mathrm{d}t.$$

If $\lambda = \lambda_k = \cdots = \lambda_{k+r-1}$ is an eigenvalue of the kernel of multiplicity r, then the Third Fredholm Theorem states that the integral equation has solutions if and only if $f(x)$ is orthogonal to all of the eigenfunctions of the homogeneous adjoint equation corresponding to $\overline{\lambda}$. These are the eigenfunctions of $\overline{K(t,x)}$; however, since $K(x,t)$ is Hermitian, they are the same as the eigenfunctions of $K(x,t)$. Also, by Proposition 3.2.3, λ is real. Thus, for solutions to exist in this case, it is necessary and sufficient that $\langle f, \phi_j \rangle = 0$ for all $j = k, \ldots, k+r-1$.

If all of these orthogonality conditions are satisfied, then the solution to the inhomogeneous integral equation has the form

$$\phi(x) = f(x) + \lambda_k \phi^{(p)}(x;\lambda_k) + \beta \phi^{(h)}(x;\lambda_k),$$

where $f(x) + \phi^{(p)}(x;\lambda_k)$ is a particular solution to the integral equation, β is an arbitrary constant, and $\phi^{(h)}(x;\lambda_k)$ is an arbitrary linear combination of the eigenfunctions corresponding to λ_k.

If j is an index for which $\lambda_j \neq \lambda_k$, then by forming the scalar product of $\phi(x)$ with $\phi_j(x)$, we obtain

$$\langle \phi, \phi_j \rangle = \langle f, \phi_j \rangle + \lambda_k \frac{\langle \phi, \phi_j \rangle}{\lambda_j}$$

or, equivalently,

$$\langle \phi, \phi_j \rangle = \frac{\lambda_j}{\lambda_j - \lambda_k} \langle f, \phi_j \rangle.$$

By substituting these values, we obtain the first sum in the proposed form of the solution.

The indices j for which $\lambda_j = \lambda_k$ correspond to the second sum in the given solution which consists of an arbitrary linear combination of the corresponding eigenfunctions. □

Illustrative Examples

- *Example 1*: Consider the Fredholm integral equation of the first kind

$$f(x) = \int_a^b K(x,t)\,\phi(t)\,\mathrm{d}t, \tag{3.7}$$

where $f(x)$ is continuous on the interval $[a,b]$ and the kernel $K(x,t)$ is non-vanishing, complex-valued, continuous, and Hermitian on the square $Q(a,b)$. In Example 4 of Sect. 3.3, we observed that any solution to this equation must have the specific form

$$\phi(x) = \sum_{n=1}^{N} \lambda_n \langle f, \phi_n \rangle \phi_n(x) + \Phi(x),$$

where $\Phi(x)$ is orthogonal to all N of the eigenfunctions of the separable kernel.

In this example, we assume that the Hermitian kernel $K(x,t)$ has an infinite number of eigenvalues. If we assume that the integral equation has a solution $\phi \in \mathcal{R}^2[a,b]$, then we can invoke the Hilbert–Schmidt Theorem to reformulate the integral equation as

$$f(x) = \sum_{n=1}^{\infty} \frac{\langle \phi, \phi_n \rangle}{\lambda_n} \phi_n(x).$$

If the mutually orthonormal set $\{\phi_n(x)\}_{n=1}^{\infty}$ of eigenfunctions is complete in $\mathcal{R}^2[a,b]$, then we can also represent $f(x)$ in the form

$$f(x) = \sum_{n=1}^{\infty} \langle f, \phi_n \rangle \, \phi_n(x).$$

Consequently, $\langle \phi, \phi_n \rangle = \lambda_n \langle f, \phi_n \rangle$, and

$$\phi(x) = \sum_{n=1}^{N} \lambda_n \langle f, \phi_n \rangle \, \phi_n(x).$$

If the set $\{\phi_n(x)\}_{n=1}^{\infty}$ of eigenfunctions is not complete, then there exist functions $\Phi(x) \in \mathcal{R}^2[a,b]$ for which $\langle \Phi, \phi_n \rangle = 0$ for all $n \geq 1$. But then also $\langle \Phi, f \rangle = 0$, so that the solution assumes the form

$$\phi(x) = \sum_{n=1}^{N} \lambda_n \langle f, \phi_n \rangle \, \phi_n(x) + \Phi(x).$$

Note that

$$\sum_{n=1}^{N} \lambda_n^2 |\langle f, \phi_n \rangle|^2 = \sum_{n=1}^{N} |\langle \phi, \phi_n \rangle|^2 \leq \|\phi\|_2^2$$

by Bessel's inequality. Thus, for a solution to exist, it is not only necessary for $f(x)$ to be expressible in terms of the eigenfunctions of $K(x,t)$, but it is also necessary for the Fourier coefficients of $f(x)$ to satisfy a restrictive growth condition.

In more advanced textbooks, equations of the first kind are considered in case $f(x)$ is square integrable, either in the sense of Riemann or in the sense of Lebesgue. We do not consider this possibility here.

It may be possible to solve an integral equation of the first kind even if the kernel is *not* Hermitian by using the methods in this section. If we multiply equation (3.7) by $K^*(s,x)$ and then integrate with respect to x, we obtain

$$\begin{aligned}
g(s) &= \mathbf{K}^* f(s) \\
&= \int_a^b K^*(s,x) f(x) \, \mathrm{d}x \\
&= \int_a^b K^*(s,x) \left(\int_a^b K(x,t) \, \phi(t) \, \mathrm{d}t \right) \mathrm{d}x \\
&= \int_a^b \left(\int_a^b K^*(s,x) K(x,t) \, \mathrm{d}x \right) \phi(t) \, \mathrm{d}t \\
&= \int_a^b H(s,t) \, \phi(t) \\
&= \mathbf{H}\phi(s).
\end{aligned}$$

If $f(x)$ is continuous, then so is $g(s)$. The composed kernel $H(s,t)$ is a nontrivial, continuous, Hermitian kernel. (Actually, it is normal.) Denote its eigenvalues by $\{\nu_n\}_{n=1}^{\infty}$ and its corresponding orthonormal eigenfunctions by $\{\varphi_n(s)\}_{n=1}^{\infty}$. If $\varphi_n = \nu_n \mathbf{H}\varphi_n$, then

$$\begin{aligned}\|\varphi_n\|_2^2 &= \langle \varphi_n, \varphi_n \rangle \\ &= \nu_n \langle \mathbf{H}\varphi_n, \varphi_n \rangle \\ &= \nu_n \langle \mathbf{K}^*\mathbf{K}\varphi_n, \varphi_n \rangle \\ &= \nu_n \langle \mathbf{K}\varphi_n, \mathbf{K}\varphi_n \rangle \\ &= \nu_n \|\mathbf{K}\phi_n\|_2^2.\end{aligned}$$

Consequently, the eigenvalues of $H(s,t)$ are all positive, i.e., $H(s,t)$ is a positive kernel.

If $\phi(x)$ is a solution of the integral equation (3.7), then it also is a solution of the equation $g(s) = \mathbf{H}\phi(s)$. Consequently, $\phi(s)$ can be expressed in terms of the eigenvalues and eigenfunctions of $H(s,t)$. On the one hand, we have

$$g(s) = \sum_{n=1}^{\infty} \langle g, \varphi_n \rangle \varphi_n(s).$$

On the other hand, we also have

$$\mathbf{H}\phi(s) = \sum_{n=1}^{\infty} \frac{\langle \phi, \varphi_n \rangle}{\nu_n} \varphi_n(s).$$

Consequently, $\langle \phi, \varphi_n \rangle = \nu_n \langle g, \varphi_n \rangle$, and

$$\phi(x) = \sum_{n=1}^{\infty} \langle \phi, \varphi_n \rangle \varphi_n(x) = \sum_{n=1}^{\infty} \nu_n \langle g, \varphi_n \rangle \varphi_n(x).$$

- *Example 2*: In this example, we show that if the inequality

$$\sum_{i=1}^{n} \sum_{j=1}^{n} K(x_i, x_j)\, \eta_i \bar{\eta}_j \geq 0$$

holds for arbitrary x_i, with $a = x_0 < x_1 < \cdots < x_n = b$ and arbitrary complex numbers $\eta_1, \ldots, \eta_n$, then $K(x,t)$ is a positive kernel. To do so, we review the concept of the double integral of a continuous function.

Let $\Delta x = \Delta t = (b-a)/n$, and define $x_i = t_i = a + \mathrm{i}\Delta x$. With this partition of the interval $[a,b]$, the square $Q(a,b)$ can be partitioned into subsquares $Q_{ij} = \{(x,t)\colon x_{i-1} \leq x \leq x_i \quad \text{and} \quad t_{i-1} \leq t \leq t_i\}$, each of which has area equal to $\Delta x \Delta t$.

If $F(x,t)$ is real- or complex-valued and continuous on the square $Q(a,b)$, then we define the double integral of $F(x,t)$ to be

$$\iint_{Q(a,b)} F(x,t)\,\mathrm{d}x\,\mathrm{d}t = \lim_{n\to\infty}\left(\sum_{i=1}^{n}\sum_{j=1}^{n} F(x_i^*,t_j^*)\,\Delta x\,\Delta t\right),$$

where the points $(x_i^*,t_j^*) \in Q_{ij}$ can be chosen arbitrarily.

The inner product $\langle \mathbf{K}\phi,\phi\rangle$ can be written as a double integral. Indeed, we have

$$\begin{aligned}\langle \mathbf{K}\phi,\phi\rangle &= \int_a^b \left(\int_a^b K(x,t)\,\phi(t)\,\mathrm{d}t\right)\overline{\phi(x)}\,\mathrm{d}x\\ &= \int_a^b\int_a^b K(x,t)\,\overline{\phi(x)}\,\phi(t)\,\mathrm{d}x\,\mathrm{d}t\\ &= \lim_{n\to\infty}\left(\sum_{i=1}^{n}\sum_{j=1}^{n} K(x_i^*,t_j^*)\,\overline{\phi(x_i^*)}\,\phi(t_j^*)\,\Delta x\,\Delta t\right).\end{aligned}$$

If we choose $x_i^* = x_i$ and $\eta_i = \overline{\phi(x_i^*)}$ for $i = 1,\ldots,n$ in the assumed condition, then we can conclude that $\langle \mathbf{K}\phi,\phi\rangle \geq 0$ as required.

The converse to this assertion also holds. A proof by contradiction can be constructed, although we shall not do so here.

Section 3.4 Exercises

1. Suppose that $f(x)$ is defined and continuous on the interval $[-\pi,+\pi]$, and that its definition is extended to be continuous and periodic on the real line $\mathbb{R}$ so that $f(x+2\pi) = f(x)$. Suppose further that its derivative $f'(x)$ is piecewise continuous on $\mathbb{R}$. If $f(x)$ is even, then it has a Fourier cosine development of the form

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(nx)$$

which converges to $f(x)$ on $\mathbb{R}$. Explain why

$$a_n = 2\int_0^{\pi} f(x)\cos(nx)\,\mathrm{d}x$$

for all $n \geq 0$.

2. Assuming that $f(x)$ has the properties as stated in Exercise 1, define the real symmetric kernel

$$L(x,t) = \frac{f(x+t)+f(x-t)}{2}$$

on the square $Q(-\pi,+\pi)$.

(a) Show that $L(x,t)$ has the double Fourier cosine representation

$$L(x,t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(nx)\cos(nt)$$

on the square $Q(a,b)$.
(b) Compute the norm $\|L\|_2$.
(c) Determine the eigenvalues and eigenfunctions of $L(x,t)$.
(d) Compute the iterated kernels $L_m(x,t)$ of $L(x,t)$, their traces A_m, and the corresponding resolvent kernel $R_L(x,t;\lambda)$.

3. Assuming that $f(x)$ has the properties as stated in Exercise 1, define the real symmetric kernel

$$M(x,t) = \frac{f(x+t) - f(x-t)}{2}$$

on the square $Q(-\pi,+\pi)$.

(a) Show that $M(x,t)$ has the double Fourier sine representation

$$M(x,t) = \sum_{n=1}^{\infty} a_n \sin(nx)\sin(nt)$$

on the square $Q(a,b)$.
(b) Compute the norm $\|M\|_2$.
(c) Determine the eigenvalues and eigenfunctions of $M(x,t)$.
(d) Compute the iterated kernels $M_m(x,t)$ of $M(x,t)$, their traces A_m, and the corresponding resolvent kernel $R_M(x,t;\lambda)$.

4. Let $f(x) = |x|$ on the interval $[-\pi,+\pi]$ and extend its definition to $\mathbb{R}$ to be periodic so that $f(x+2\pi) = f(x)$.

(a) Show that $f(x)$ has the Fourier cosine development

$$f(x) = \frac{\pi}{2} - \frac{4}{\pi}\sum_{n=1}^{\infty} \frac{1}{(2n-1)^2}\cos[(2n-1)x].$$

(b) Define the real symmetric kernel

$$L(x,t) = \frac{|x+t| + |x-t|}{2}$$

on the square $Q(-\pi,+\pi)$. Show that $L(x,x) = L(x,-x) = |x|$ and that $L(x,t) = L(-x,t) = L(x,-t) = L(-x,-t)$. Deduce from these equations that $L(x,t)$ has an alternate form $L(x,t) = \max\{|x|,|t|\}$. Use this information to sketch the surface $z = L(x,t)$ above the square $Q(-\pi,+\pi)$. (This surface resembles an inverted pyramid; the horizontal cross sections are squares.)

(c) Determine the infinite bilinear expansion of $L(x,t)$, and discuss its convergence properties on $Q(-\pi,+\pi)$.
(d) What are the eigenvalues and eigenfunctions of $L(x,t)$?
(e) Compute the iterated kernels $L_m(x,t)$ of $L(x,t)$, their traces A_m, and the corresponding resolvent kernel $R_L(x,t;\lambda)$.

5. Let $f(x) = |x|$ on the interval $[-\pi,+\pi]$ and extend its definition to $\mathbb{R}$ to be periodic so that $f(x+2\pi) = f(x)$.

(a) Define the real symmetric kernel

$$M(x,t) = \frac{|x+t| - |x-t|}{2}$$

on the square $Q(-\pi,+\pi)$. Show that $M(x,x) = |x| = -M(x,-x)$, that $M(x,t) = -M(-x,t) = -M(x,-t) = +M(-x,-t)$, and that $M(0,t) = M(x,0) = 0$. Deduce from these observations that $M(x,t)$ has the alternate form

$$M(x,t) = \begin{cases} +\min\{|x|,|t|\} & \text{if } xt \geq 0 \\ -\min\{|x|,|t|\} & \text{if } xt \leq 0. \end{cases}$$

Use this information to sketch the surface $z = M(x,t)$ above the square $Q(-\pi,+\pi)$.
(b) Determine the infinite bilinear expansion of $M(x,t)$ and discuss its convergence properties on $Q(-\pi,+\pi)$.
(c) What are the eigenvalues and eigenfunctions of $M(x,t)$?
(d) Compute the iterated kernels $M_m(x,t)$ of $M(x,t)$, their traces A_m, and the corresponding resolvent kernel $R_M(x,t;\lambda)$.

6. For $-1 < r < 1$ and $0 \leq \theta \leq 2\pi$, define

$$P(\theta) = \sum_{n=-\infty}^{n=+\infty} r^{|n|}\, \mathrm{e}^{\mathrm{i}n\theta}.$$

(a) Use the formula for the sum of a geometric series to show that

$$P(\theta) = \frac{1-r^2}{1-2r\cos\theta + r^2}.$$

(b) Combine the terms of $P(\theta)$ to show also that

$$P(\theta) = 1 + \sum_{n=1}^{\infty} 2r^n \cos(n\theta).$$

(c) Show that

$$\int_0^{2\pi} \cos^2(nx)\,\mathrm{d}x = \int_0^{2\pi} \sin^2(nx)\,\mathrm{d}x = \pi,$$

and use these facts to show that the set

$$\left\{\phi_0(x) = \frac{1}{\sqrt{2\pi}}\right\} \cup \left\{\phi_n(x) = \frac{\cos(nx)}{\sqrt{\pi}},\ \psi_n(x) = \frac{\sin(nx)}{\sqrt{\pi}}\right\}_{n=1}^{\infty}$$

is a mutually orthonormal system on the interval $[0, 2\pi]$.

(d) For $0 \le x \le 2\pi$ and $0 \le t \le 2\pi$, define the real symmetric kernel

$$D(x,t) = \frac{1}{2\pi}\left(1 + \sum_{n=1}^{\infty} 2r^n \cos(n(x-t))\right).$$

By using a trigonometric identity, convert $D(x,t)$ into the form

$$D(x,t) = \frac{\phi_0(x)\phi_0(t)}{\sqrt{2\pi}} + \sum_{n=1}^{\infty}\left[\frac{\phi_n(x)\phi_n(t)}{r^{-n}} + \frac{\psi_n(x)\psi_n(t)}{r^{-n}}\right],$$

displaying its eigenvalues and eigenfunctions.

(e) Compute the iterated kernels $D_n(x,t)$, their traces, and the resolvent kernel $R_D(x,t;\lambda)$.

7. If $G(x,t)$ is an arbitrary real-valued continuous kernel defined on the square $Q(a,b)$, then the composed kernels

$$G_L(x,t) = \int_a^b G(x,s)\,G(t,s)\,\mathrm{d}s$$

and

$$G_R(x,t) = \int_a^b G(s,x)\,G(s,t)\,\mathrm{d}s$$

are real-valued, continuous symmetric kernels. Let $\mathbf{G}_L$ and $\mathbf{G}_R$ denote the Fredholm operators that correspond to $G_L(x,t)$ and $G_R(x,t)$.

(a) Prove that $\langle \mathbf{G}_L\phi, \phi\rangle \ge 0$ and that $\langle \mathbf{G}_R\phi, \phi\rangle \ge 0$, showing that the eigenvalues of $G_L(x,t)$ and $G_R(x,t)$ are positive.

(b) Are the eigenvalues of $G_L(x,t)$ equal to the eigenvalues of $G_R(x,t)$?

(c) What can you say about their eigenfunctions? Are they related?

8. In Proposition 3.4.1(1), it was established that all of the eigenvalues of a nonvanishing, complex-valued, continuous, Hermitian, positive kernel are positive numbers. Does the converse hold?

9. The kernel $K(x,t)$ is called *negative* if $\langle \mathbf{K}\phi, \phi\rangle \le 0$ for all $\phi \in \mathcal{R}^2[a,b]$. Explain why all of the eigenvalues of $K(x,t)$ are negative numbers.

10. Explain why Mercer's theorem holds if $K(x,t)$ is a negative kernel.
11. Explain why Mercer's theorem holds if

    (a) All but finitely many of the eigenvalues of $K(x,t)$ are positive.
    (b) All but finitely many of the eigenvalues of $K(x,t)$ are negative.

12. Suppose that $K(x,t)$ is a nonvanishing, complex-valued, continuous, and Hermitian, positive kernel. By Proposition 3.2.1(3), $K(x,x)$ is real, and by Proposition 3.4.1(2), $K(x,x) \geq 0$. By means of an example, show that it is possible for $K(x,t) < 0$ if $x \neq t$.
13. Invent a kernel for which the traces $A_{2m+1} = 0$ for all $m \geq 0$.
14. Let

$$K(x,t) = \begin{cases} 1 & \text{if } 0 \leq t \leq x \leq 1 \\ 0 & \text{if } 0 \leq x < t \leq 1. \end{cases}$$

    Clearly, $K(x,t)$ is neither continuous nor Hermitian. Determine $K^*(x,t)$ and then show that the composed kernel $H = K^*K$ is both continuous and Hermitian.
15. Suppose that a sequence of complex-valued, nonvanishing, continuous, Hermitian, positive kernels converges uniformly to a limit kernel $L(x,t)$ on the square $Q(a,b)$. Does $L(x,t)$ also have all of these properties as well?

3.5 Numerical Methods

In this section, we discuss some numerical techniques for computing valuable approximations to the eigenvalues and eigenfunctions of a Hermitian kernel.

3.5.1 The Method of Traces

Lemma 3.2.1 (The Trace Lemma) established some basic facts about the traces of a kernel and its iterates. In this subsection, we show how to obtain useful inequalities to estimate the smallest eigenvalue of a kernel.

Proposition 3.5.1. *Let $K(x,t)$ be a complex-valued, nonvanishing, continuous Hermitian kernel, and let A_m denote the mth trace of $K(x,t)$. Then*

1. For every $m \geq 1$, we have

$$|\lambda_1| \leq \sqrt{\frac{A_{2m}}{A_{2m+2}}}.$$

2. *The magnitude of the smallest eigenvalue of $K(x,t)$ is given by*

$$|\lambda_1| = \lim_{m\to\infty} \sqrt{\frac{A_{2m}}{A_{2m+2}}}.$$

3. *For every $m \geq 1$, we have*

$$\sqrt[2m]{\frac{g}{A_{2m}}} \leq |\lambda_1|,$$

where g is the sum of the geometric multiplicities of $+\lambda_1$ and $-\lambda_1$.

4. *The magnitude of the smallest eigenvalue of $K(x,t)$ is given by*

$$|\lambda_1| = \lim_{m\to\infty} \frac{1}{\sqrt[2m]{A_{2m}}}.$$

Proof. Let g^+ and g^- denote the geometric multiplicities of the eigenvalues $+\lambda_1$ and $-\lambda_1$, respectively, and let $g = g^+ + g^-$. (The geometric multiplicity is equal to the algebraic multiplicity.) If $K(x,t)$ has more than g eigenvalues, then for $m \geq 1$, the formula for the trace A_{2m} can be rearranged in the form

$$\begin{aligned} A_{2m} &= \sum_{n=1}^{\infty} \frac{1}{\lambda_n^{2m}} \\ &= \frac{g}{\lambda_1^{2m}} \left[1 + \frac{1}{g} \sum_{n=g+1}^{\infty} \left(\frac{\lambda_1}{\lambda_n} \right)^{2m} \right] \\ &= \frac{g}{\lambda_1^{2m}} \left(1 + \delta_m \right). \end{aligned}$$

1. Since $(\lambda_1/\lambda_n)^2 < 1$ for all $n \geq g+1$, it is clear that $\{\delta_m\}_{m=1}^{\infty}$ is a strictly decreasing sequence of positive numbers which necessarily converges to a limit $\delta \geq 0$. (Actually, $\delta = 0$.) The first claim follows directly from the observation that

$$\lambda_1^2 \leq \frac{1+\delta_m}{1+\delta_{m+1}} = \frac{A_{2m}}{A_{2m+2}}.$$

2. Consequently,

$$\lambda_1^2 = \lambda_1^2 \frac{1+\delta}{1+\delta} = \lambda_1^2 \lim_{m\to\infty} \frac{1+\delta_m}{1+\delta_{m+1}} = \lim_{m\to\infty} \frac{A_{2m}}{A_{2m+2}}.$$

3. It also follows from the rearrangement above that

$$\sqrt[2m]{\frac{g}{A_{2m}}} \leq \sqrt[2m]{\frac{g(1+\delta_m)}{A_{2m}}} = |\lambda_1| \leq \sqrt[2m]{\frac{g(1+\delta_1)}{A_{2m}}}.$$

4. Since

$$\lim_{m\to\infty} \sqrt[2m]{g} = \lim_{m\to\infty} \sqrt[2m]{g(1+\delta_1)} = 1,$$

the result follows from an application of the Squeezing Theorem. □

Unfortunately, the Method of Traces does not produce an eigenfunction that corresponds to the eigenvalue of smallest modulus.

3.5.2 The Method of Rayleigh and Ritz

In Theorem 3.2.1, it was established that

$$\frac{1}{|\lambda_1|} = \max_{\phi} |\langle \mathbf{K}\phi, \phi\rangle|,$$

where λ_1 is the eigenvalue of smallest modulus of the nonvanishing, complex-valued, continuous Hermitian kernel $K(x,t)$, and the maximum is taken over all $\phi \in \mathcal{R}^2[a,b]$ for which $\|\phi\|_2 = 1$. Furthermore, it was shown that the maximum value is attained when $\phi(x)$ is an eigenfunction that corresponds to λ_1. Thus, if $\psi(x)$ is not an eigenfunction of the kernel, then $|\langle \mathbf{K}\psi, \psi\rangle|$ is merely a lower bound for $1/|\lambda_1|$ of questionable value.

The Method of Rayleigh and Ritz is an effective technique for producing a relatively accurate lower bound for $1/|\lambda_1|$ (or equivalently, an upper bound for $|\lambda_1|$), as well as a corresponding eigenfunction.

Let $\Psi = \{\psi_n(x)\}_{n=1}^{\infty}$ be a complete, linearly independent set of continuous functions defined on the interval $[a,b]$ with $\|\psi_n\|_2 = 1$ for all $n \geq 1$, and let

$$\psi(x) = \sum_{n=1}^{N} a_n \psi_n(x)$$

be an arbitrary linear combination constructed from the elements of Ψ. If Ψ is not an orthonormal set, then

$$\|\psi\|_2^2 = \sum_{m=1}^{N}\sum_{n=1}^{N} \langle \psi_m, \psi_n\rangle a_m \bar{a}_n,$$

whereas if Ψ is an orthonormal set, then

$$\|\psi\|_2^2 = \sum_{n=1}^{N} |a_n|^2.$$

Also,

$$\langle \mathbf{K}\psi, \psi\rangle = \sum_{m=1}^{N}\sum_{n=1}^{N} K_{mn} a_m \bar{a}_n,$$

where

$$K_{mn} = \int_a^b \int_a^b K(x,t)\, \psi_m(t)\, \overline{\psi_n(x)}\, \mathrm{d}t\, \mathrm{d}x.$$

The Method of Rayleigh and Ritz employs these finite linear combinations $\psi(x)$ in a clever way in order to determine a lower bound for $1/|\lambda_1|$. Essentially, if the quadratic form $|\langle \mathbf{K}\psi, \psi\rangle|$ is maximized over all such $\psi(x)$ subject to the constraint that $\|\psi\|_2 = 1$, then the attained maximum value serves as an effective lower bound for $1/|\lambda_1|$, and the coefficients a_n that yield the attained maximum value can be used to construct an approximation to an eigenfunction.

In general, several complicating factors may be present:

- The eigenvalue λ_1 of smallest modulus may be positive or negative.
- The set Ψ may consist of complex-valued functions.
- The number N of linearly independent functions composing $\psi(x)$ may be large.
- The coefficients a_n may be complex numbers.
- There may be more than one eigenfunction corresponding to λ_1.

Accordingly, we make the following simplifying assumptions:

- The kernel $K(x,t)$ is real symmetric.
- The smallest eigenvalue of $K(x,t)$ is positive.
- The system Ψ is a mutually orthonormal complete system that consists of real-valued functions.
- The coefficients a_n are real numbers.
- There is only one eigenfunction corresponding to λ_1.

Under these assumptions, we can rephrase our problem in the following way:

Maximize the real quadratic form

$$F(a_1,\ldots,a_N) = \sum_{m=1}^{N} \sum_{n=1}^{N} K_{mn}\, a_m\, a_n$$

subject to the constraint

$$G(a_1,\ldots,a_N) = \sum_{n=1}^{N} a_n^2 - 1 = 0.$$

The Method of Lagrange Multipliers is ideally suited to address this problem. Since F and G are continuously differentiable on $\mathbb{R}^N$ and ∇G does not vanish on the constraint surface, the extreme values of F occur among the solutions of the gradient equation $\nabla F = \mu\, \nabla G$ for some auxiliary parameter μ (the multiplier). The fact that the coefficient matrix $\tilde{\mathbf{K}} = (K_{mn})$ is real symmetric allows for the reformulation of the gradient equation as the linear system of N coordinate equations

$$\sum_{n=1}^{N} K_{mn}\, a_n = \mu\, a_m \quad (m = 1,\ldots,N)$$

This system can be recast in matrix form as

$$\begin{pmatrix} K_{11}-\mu & K_{12} & \cdots & K_{1N} \\ K_{21} & K_{22}-\mu & \cdots & K_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ K_{N1} & K_{N2} & \cdots & K_{NN}-\mu \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

or more elegantly as $(\tilde{\mathbf{K}}-\mu\,\mathbf{I})\mathbf{a}=\mathbf{0}$. For this homogeneous system to have at least one nontrivial solution, it is necessary for the determinant $\det(\tilde{\mathbf{K}}-\mu\,\mathbf{I})$ (the characteristic polynomial of the matrix $\tilde{\mathbf{K}}$ of degree N) to vanish. Since $\tilde{\mathbf{K}}$ is real symmetric, all of the zeroes of the determinant are real. Let μ_N^* denote the largest positive zero of the equation $\det(\tilde{\mathbf{K}}-\mu\,\mathbf{I})=0$, and let $\mathbf{a}^*=(a_1^*,\ldots,a_N^*)$ denote a solution of the system corresponding to μ_N^*. Then,

$$\sum_{n=1}^{N} K_{mn}\,a_n^* = \mu_N^*\,a_m^* \quad (m=1,\ldots,N)$$

If we multiply each of these equations by a_m^* and then sum them over the index m, we obtain

$$F(a_1^*,\ldots,a_N^*)=\mu_N^*.$$

Thus, the multiplier μ_N^* is actually the maximum value of $F(a_1,\ldots,a_N)$ subject to the imposed constraint, and this maximum value is attained at each of the corresponding solutions of the linear system.

Since we assumed for purposes of argumentative simplicity that the smallest eigenvalue of $K(x,t)$ is positive, it follows from the comments at the beginning of this subsection that $1/\lambda_1 \geq \mu_N^*$ or, equivalently, that $1/\mu_N^* \geq \lambda_1$, i.e., that $1/\mu_N^*$ is an approximation (in excess) to λ_1. Since we also assumed that there is only one eigenfunction corresponding to λ_1, the function

$$\psi_N^*(x)=\sum_{n=1}^{N} a_n^*\,\psi_n(x)$$

is an approximation to the eigenfunction $\phi_1(x)$ corresponding to λ_1.

It is clear that the sequence $\{\mu_N^*\}$ increases with N. Thus, the decreasing sequence $1/\mu_N^*$ approximates λ_1 with increasing accuracy as N increases. It can be shown that $1/\mu_N^* \to \lambda_1$ as $N\to\infty$, but we shall not do so here.

Illustrative Examples

- *Example 1*: Consider the kernel $K(x,t) = e^{xt}$ defined on the square $Q(0,1)$.

  In Exercise 3 of Sect. 3.2, it was shown that $A_1 = 1.462651$, $A_2 = 1.841935$, and $A_3 = 2.478170$, $A_4 = 3.351554$ and that all of the traces of $K(x,t)$ are positive.

  By Proposition 3.5.1(1), we immediately have the *upper estimate*

$$|\lambda_1| \leq \sqrt{\frac{A_2}{A_4}} < 0.741334.$$

  If g_1 denotes the sum of the geometric multiplicities of $+\lambda_1$ and $-\lambda_1$, then

$$A_4 = \sum_{n=1}^{\infty} \frac{1}{\lambda_n^4} \geq \sum_{n=1}^{g_1} \frac{1}{\lambda_n^4} = \frac{g_1}{\lambda_1^4}.$$

  Hence,

$$1 \leq g_1 \leq \lambda_1^4 A_4 < (0.741334)^4 (3.351554) < 1.012282,$$

  i.e., $g_1 = 1$. Thus, there is exactly one eigenfunction corresponding to λ_1.

  By Proposition 3.5.1(3), we now have the *lower estimate*

$$0.739074 < \sqrt[4]{\frac{1}{A_4}} \leq |\lambda_1|.$$

  By employing these estimates, we can provide a lower estimate on $|\lambda_2|$. If g_2 denotes the sum of the geometric multiplicities of $+\lambda_2$ and $-\lambda_2$, then

$$\frac{1}{\lambda_2^2} \leq \frac{g_2}{\lambda_2^2} \leq A_2 - \frac{1}{\lambda_1^2} < 0.022353,$$

  from which we conclude that $|\lambda_2| > 6.688536$.

- *Example 2*: In practice, when one applies the Method of Rayleigh and Ritz, a judicious choice for the mutually orthonormal system usually leads to accurate approximations. On the other hand, a poor choice could lead to decidedly disappointing results.

  To illustrate this point, consider the system $\Psi = \{\psi_n(x)\}_{n=1}^{\infty}$, where $\psi_n(x) = \sqrt{2}\sin(n\pi x)$, which is defined on the interval $[0,1]$. Note that $\psi_n(0) = \psi_n(1) = 0$ for all $n \geq 1$. If we set

$$\psi(x) = \sum_{n=1}^{N} a_n \psi_n(x),$$

  then $\psi(0) = \psi(1) = 0$, regardless of the choice of N.

If $\phi_1(x)$ is an eigenfunction corresponding to the smallest eigenvalue of some kernel and it were known beforehand that $\phi_1(0) = \phi_1(1) = 0$, then choosing Ψ would probably lead to an acceptable approximation to $\phi_1(x)$. On the other hand, if it were known in advance that $\phi_1(0) = 1$ and $\phi_1(1) = 2$, then choosing Ψ would most likely lead to an unacceptable approximation for $\phi_1(x)$, especially near the endpoints of the interval.

In general, in the absence of any a priori information about $\phi_1(x)$, it is advisable to choose a mutually orthonormal system with the greatest adaptability over the entire interval of definition. Systems of orthonormal polynomials are easy to use and lead to accurate approximations.

- *Example 3*: In Example 1 above, we used the Method of Traces to show that the smallest eigenvalue λ_1 of the kernel $K(x,t) = \mathrm{e}^{xt}$ satisfies the inequality

$$0.739074 \leq \lambda_1 \leq 0.741334,$$

which is actually quite good, considering the fact that it was computed in terms of only A_2 and A_4.

In this example, we apply the Method of Rayleigh and Ritz to improve this estimate. Our application also leads to an excellent approximation to the corresponding eigenfunction $\phi_1(x)$, where

$$\phi_1(x) = \lambda_1 \int_0^1 \mathrm{e}^{xt}\, \phi_1(t)\, \mathrm{d}t.$$

In order to employ their method, we must choose a mutually orthonormal system of functions defined on the interval $[0,1]$. In view of Example 2 above, we choose to construct an orthonormal system of polynomials. An application of the Gram–Schmidt Orthonormalization Procedure to the set $\{x^n\}_{n=0}^{\infty}$ of polynomials produces the requisite system $\Psi = \{\psi_n(x)\}_{n=1}^{\infty}$. The polynomials $\psi_n(x)$ are pairwise orthogonal and $\|\psi_n\|_2 = 1$ for all $n \geq 1$ as required.

The first four elements of Ψ are

$$\begin{aligned}
\psi_1(x) &= 1, \\
\psi_2(x) &= \sqrt{3}\,(2x-1), \\
\psi_3(x) &= \sqrt{5}\,(6x^2-6x+1),
\end{aligned}$$

and

$$\psi_4(x) = \sqrt{7}\,(20x^3-30x^2+12x-1).$$

The approximation $\psi(x)$ to $\phi_1(x)$ determined below will be written as a linear combination of these four polynomials.

Now that the orthonormal system Ψ has been chosen, the next step is to compute the entries of the matrix $\tilde{\mathbf{K}} = (K_{mn})$, where

$$K_{mn} = \int_0^1 \int_0^1 \mathrm{e}^{xt}\, \psi_m(x)\, \psi_n(t)\, \mathrm{d}x\, \mathrm{d}t.$$

After these computations have been completed, we find that

$$\tilde{\mathbf{K}} = \begin{pmatrix} 1.317902 & 0.205527 & 0.018360 & 0.001189 \\ 0.205527 & 0.138880 & 0.017096 & 0.001291 \\ 0.018360 & 0.017096 & 0.005717 & 0.000625 \\ 0.001189 & 0.001291 & 0.000625 & 0.000148 \end{pmatrix}.$$

Note that the matrix $\tilde{\mathbf{K}}$ is real symmetric, so that all of its eigenvalues are real. These eigenvalues are the solutions of the polynomial equation $\det(\tilde{\mathbf{K}} - \mu\, \mathbf{I}) = 0$, which we determine to be $1.353030, 0.105983, 0.003560$, and 0.000075, accurate to six decimal places. If we denote the largest of these eigenvalues by μ_4^*, then its reciprocal serves as an upper bound for λ_1, i.e., $\lambda_1 \le 1/\mu_4^* = 1/1.353030 = 0.739081$.

By combining this estimate with the lower estimate in Example 1, we obtain the improved estimate $0.739074 \le \lambda_1 \le 0.739081$.

Finally, by solving the homogeneous linear system $\left(\tilde{\mathbf{K}} - \mu_4^* \mathbf{I}\right) \mathbf{a} = \mathbf{0}$, we obtain the coefficients

$$\mathbf{a}^* = \langle a_1^*, a_2^*, a_3^*, a_4^* \rangle = \langle -0.985817, -0.167096, -0.015555, -0.001033 \rangle.$$

Hence, the approximation $\psi(x)$ to the eigenfunction $\phi_1(x)$ is given by

$$\begin{aligned} \psi(x) &= a_1^*\, \psi_1(x) + a_2^*\, \psi_1(x) + a_3^*\, \psi_3(x) + a_4^*\, \psi_4(x) \\ &= -0.728445 - 0.402959\, x - 0.126654\, x^2 - 0.054694\, x^3. \end{aligned}$$

To illustrate the accuracy of our calculations, we estimate the residual

$$\eta(x) = \psi(x) - 0.739081 \int_0^1 \mathrm{e}^{xt}\, \psi(t)\, \mathrm{d}t$$

in two ways: We find that $\|\eta\|_\infty = 0.000166$ and that $\|\eta\|_2 = 0.000053$.

Section 3.5 Exercises

1. The results discussed in this section illustrate two methods that can be used to approximate the magnitude of the smallest eigenvalue $|\lambda_1|$ of a kernel and a corresponding eigenfunction. If λ_1 is known and there is exactly one real-valued eigenfunction $\phi_1(x)$ corresponding to it, then the truncated kernel

$$\Delta_2(x,t) = K(x,t) - \frac{\phi_1(x)\,\phi_1(t)}{\lambda_1}$$

can be constructed.

   (a) Review Lemma 3.3.1 (Kernel structure). Explain why the magnitude of the smallest eigenvalue of $\Delta_2(x,t)$ is $|\lambda_2|$.
   (b) Review Theorem 3.2.1. Consider the problem of computing

$$\max_{\phi} |\langle \mathbf{D}_2\phi, \phi\rangle|$$

   over all $\phi(x)$ such that $\|\phi\|_2 = 1$ and $\langle \phi, \phi_1\rangle = 0$, where $\mathbf{D}_2$ is the Fredholm operator corresponding to the truncated kernel $\Delta_2(x,t)$. Is the maximum value attained? If so, for which ϕ?
   (c) Theoretically, can this process of truncation be repeated? If so, can you approximate $|\lambda_3|$? $|\lambda_4|$?
   (d) Practically speaking, if this process is repeated several times, what happens to the numerical accuracy of your approximations?

2. Consider the kernel $C(x,t) = \cosh(xt)$.

   (a) Follow the procedures illustrated in Example 1 to compute the traces A_1, A_2, A_3, and A_4.
   (b) Use the Method of Traces to estimate $|\lambda_1|$.
   (c) Use the Method of Rayleigh and Ritz to estimate $|\lambda_1|$ and $\phi_1(x)$.

3. Consider the kernel $G(x,t) = \dfrac{1}{1+xt}$.

   (a) Show that $A_4 = \|G\|_2^2 = (2\pi^2 \ln 2 - 9\,\zeta(3))/6$, where $\zeta(z)$ is the Riemann zeta function. (Refer to Example 7 in Sect. 3.2 and Lemma 3.2.1(2).)
   (b) Use the results of Example 7 and part (a) to estimate $|\lambda_1|$.

4. Consider the kernel

$$J(x,t) = \begin{cases} -\sqrt{xt}\,\ln t & \text{if } 0 < x \le t \le 1 \\ -\sqrt{xt}\,\ln x & \text{if } 0 < t \le x \le 1 \\ 0 & \text{if } x = t = 0 \end{cases}$$

defined on the square $Q(0,1)$. Note that $J(x,t) \equiv 0$ on $\partial Q(0,1)$ and that $J(x,t)$ is positive on the interior of $Q(0,1)$. Note also that we defined

$$J(0,0) = \lim_{(x,t)\to(0,0)} J(x,t) = 0$$

by continuous extension.

(a) Show that

$$A_1 = \int_0^1 J(x,x)\,\mathrm{d}x = -\int_0^1 x \ln x\,\mathrm{d}x = \tfrac{1}{4}.$$

(b) Compute the iterated kernel $J_2(x,t)$.
(c) Compute $A_2 = \|J\|_2^2$ and then check your answer by computing

$$A_2 = \int_0^1 J_2(x,x)\,\mathrm{d}x.$$

(d) Compute $A_4 = \|J_2\|_2^2$.
(e) Use parts (c) and (d) to estimate the smallest eigenvalue of $J(x,t)$.

Chapter 4
Volterra Integral Equations

In this chapter, our attention is devoted to the Volterra integral equation of the second kind which assumes the form

$$\phi(x) = f(x) + \lambda \int_a^x K(x,t)\,\phi(t)\,\mathrm{d}t. \tag{4.1}$$

Volterra integral equations differ from Fredholm integral equations in that the upper limit of integration is the variable x instead of the constant b.

In our analysis, we make the following assumptions:

- The unknown function $\phi(x)$ is always assumed to be integrable in the sense of Riemann, so that the integral equation itself makes sense.
- The *free term* $f(x)$ is assumed to be complex-valued and continuous on the interval $[a,b]$.
- The complex constant $\lambda\ (\neq 0)$ is a parameter.
- The *kernel* $K(x,t)$ is assumed to be complex-valued and continuous on the triangle $T(a,b) = \{(x,t)\colon a \le x \le b, a \le t \le x\}$. Volterra kernels are always assumed to satisfy the condition $K(x,t) \equiv 0$ if $x < t$, i.e., the kernel vanishes above the diagonal of the square $Q(a,b)$.

In Sect. 4.1, we present several tools of the trade which are indispensible for the comprehension of the material in the chapter.

In Sect. 4.2, we revisit the Method of Successive Approximation which can be applied to produce a solution to the integral equation in the form of a Maclaurin series with an infinite radius of convergence in the variable λ.

In Sect. 4.3, we show that the Laplace transformation can be used to solve the integral equation when the kernel assumes the form $K(x,t) = k(x-t)$.

In Sect. 4.4, we treat Volterra integral equations of the first kind.

S.M. Zemyan, *The Classical Theory of Integral Equations: A Concise Treatment*,
DOI 10.1007/978-0-8176-8349-8_4, © Springer Science+Business Media, LLC 2012

4.1 Tools of the Trade

In this chapter, the reader should be familiar with the following topics:

- *The Laplace transform*: Not only is the Laplace transform method an extremely useful tool for solving linear ordinary differential equations, but it is also quite valuable for solving linear Volterra integral equations of a certain type.

  Let $f(t)$ be defined on the interval $[0,\infty)$. The *Laplace transform* of $f(t)$ is given by the improper integral

$$\mathcal{L}\{f(t)\} = \int_0^\infty \mathrm{e}^{-st} f(t)\,\mathrm{d}t = \lim_{A\to\infty} \int_0^A \mathrm{e}^{-st} f(t)\,\mathrm{d}t,$$

  provided that the integral exists. The integral will exist if $f(t)$ is piecewise continuous on the interval $[0,A]$ for any A, and is of exponential order. (Recall that a function $f(t)$ is *piecewise continuous* on an interval $[0,A]$ if it is continuous there except for a finite number of jump discontinuities. Also, a function $f(t)$ is of *exponential order* if there exist constants a, c, and m such that $|f(t)| \le c\,\mathrm{e}^{at}$ for all $t \ge m$.)

  To illustrate the definition, elementary integrations show that

$$\mathcal{L}\{t^n\} = \frac{n!}{s^{n+1}}, \quad \mathcal{L}\{\mathrm{e}^{at}\} = \frac{1}{s-a}, \quad \text{and} \quad \mathcal{L}\{\sin(at)\} = \frac{a}{s^2+a^2}.$$

  The Laplace transform of the derivatives $f^{(n)}(t)$ of $f(t)$ can be expressed in terms of the Laplace transform of $f(t)$. The precise formula is

$$\mathcal{L}\left\{f^{(n)}(t)\right\} = s^n\,\mathcal{L}\{f(t)\} - \sum_{m=0}^{n-1} f^{(m)}(0)\,s^{n-1-m}.$$

  This fact is the reason why the Laplace transform can be used to solve linear ordinary differential equations with constant coefficients.

  If $f(t)$ and $g(t)$ are integrable on the interval $[0,\infty)$, then the *convolution product* of $f(t)$ and $g(t)$ is defined by the integral

$$(f * g)(t) = \int_0^t f(t-u)\,g(u)\,\mathrm{d}u.$$

  If $\mathcal{L}\{f(t)\}$ and $\mathcal{L}\{g(t)\}$ exist, then $\mathcal{L}\{(f*g)(t)\} = \mathcal{L}\{f(t)\}\,\mathcal{L}\{g(t)\}$.

- *Differentiation of a multivariable integral*: Let $F(x,t)$ be a complex-valued continuous function on the square $Q(a,b)$. If $F(x,t)$ is a continuously differentiable function of x, then

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\int_a^x F(x,t)\,\mathrm{d}t\right) = F(x,x) + \int_a^x \frac{\partial F}{\partial x}(x,t)\,\mathrm{d}t.$$

4.2 The Method of Successive Approximation

Consider the integral equation

$$\phi(x) = f(x) + \lambda \int_a^x K(x,t)\,\phi(t)\,\mathrm{d}t$$

in which the free term $f(x)$ is complex-valued and continuous on the interval $[a,b]$ and the kernel $K(x,t)$ is complex-valued and continuous on the triangle $T(a,b) = \{(x,t)\colon a \le x \le b, a \le t \le x\}$. It is always assumed that Volterra kernels satisfy the condition $K(x,t) \equiv 0$ if $x < t$, i.e., the kernel vanishes above the diagonal of the square $Q(a,b)$. If $\lambda = 0$, then $\phi(x) = f(x)$ is the unique solution to the integral equation.

If $|\lambda|$ is small, then it is reasonable to suggest that $\phi(x) \approx f(x)$, i.e., that the free term is a serviceable zeroth-order approximation $\phi_0(x)$ to the solution of the equation, provided that a solution exists.

A first-order approximation $\phi_1(x)$ to $\phi(x)$ is furnished by replacing $\phi(t)$ by $\phi_0(t) = f(t)$ in the integrand to obtain

$$\phi_1(x) = f(x) + \lambda \int_a^x K(x,t)\,\phi_0(t)\,\mathrm{d}t.$$

If the integral

$$\int_a^x K(x,t)\,\phi_0(t)\,\mathrm{d}t = 0,$$

then $\phi_1(x) = f(x) = \phi_0(x)$, and the iterative process terminates here. To show that this eventuality can actually occur, consider the equation

$$\phi(x) = x + \lambda \int_0^x (2x - 3t)\,\phi(t)\,\mathrm{d}t.$$

If we choose $\phi_0(x) = f(x) = x$, then

$$\int_0^x (2x - 3t)\,t\,\mathrm{d}t = \int_0^x 2xt - 3t^2\,\mathrm{d}t = xt^2 - t^3\Big|_0^x = 0.$$

Hence, $\phi_1(x) = f(x) = x = \phi(x)$ for every value of λ.

If $\phi_1(x) \neq \phi_0(x) = f(x)$, then a substitution of $\phi_1(t)$ into the integrand yields the second-order approximation

$$\phi_2(x) = f(x) + \lambda \int_a^x K(x,t)\,\phi_1(t)\,\mathrm{d}t.$$

Continuing in this manner, the nth-order approximation

$$\phi_n(x) = f(x) + \lambda \int_a^x K(x,t)\,\phi_{n-1}(t)\,\mathrm{d}t$$

is obtained, assuming that the integral does not vanish at any step. If the integral does vanish, then $\phi_n(x) = f(x) = \phi_0(x)$, and the iterative process fails.

Each approximant $\{\phi_n(x)\}$ has an alternate form. If we substitute the first-order approximation into the second-order approximation, then we obtain

$$\begin{aligned}
\phi_2(x) &= f(x) + \lambda \int_a^x K(x,t) \left[f(t) + \lambda \int_a^t K(t,s)\,f(s)\,\mathrm{d}s \right] \mathrm{d}t \\
&= f(x) + \lambda \int_a^x K(x,t)\,f(t)\,\mathrm{d}t + \lambda^2 \int_a^x \int_a^t K(x,t)\,K(t,s)\,f(s)\,\mathrm{d}s\,\mathrm{d}t \\
&= f(x) + \lambda \int_a^x K(x,t)\,f(t)\,\mathrm{d}t + \lambda^2 \int_a^x K_2(x,t)\,f(t)\,\mathrm{d}t,
\end{aligned}$$

where we have set

$$K_2(x,t) = \int_t^x K(x,s)\,K(s,t)\,\mathrm{d}s.$$

Note that the iterated kernel $K_2(x,t) \equiv 0$ if $x < t$. This is due to the fact that $K(x,s) \equiv 0$ whenever $x < s$ and $K(s,t) \equiv 0$ whenever $s < t$. Since these s-intervals overlap whenever $x < t$, it follows that the integrand differs from zero only when $t \le s \le x$.

Additional iterations lead to the general form

$$\phi_n(x) = f(x) + \sum_{m=1}^{n} \lambda^m \left(\int_a^x K_m(x,t)\,f(t)\,\mathrm{d}t \right), \tag{4.2}$$

where for each $m = 1, 2, \ldots$, we have set

$$K_m(x,t) = \int_t^x K_{m-1}(x,s)\,K(s,t)\,\mathrm{d}s.$$

By induction, each iterated kernel satisfies the condition $K_m(x,t) \equiv 0$ if $x < t$.

The sequence $\{\phi_n(x)\}$ of continuous approximants converges absolutely and uniformly on the interval $[a,b]$. Since we have assumed that $K(x,t)$ is continuous on the closed triangle $T(a,b)$ and that $K(x,t) \equiv 0$ if $x < t$, there exists an M such that $|K(x,t)| \le M$ on the square $Q(a,b)$. Hence,

$$|K_2(x,t)| \le M^2 \left| \int_t^x \mathrm{d}s \right| = M^2\,(x-t) \le M^2\,(b-a)$$

for $t \leq x$. We have already observed that $K_2(x,t) \equiv 0$ if $x < t$. A short inductive argument shows that the inequalities

$$|K_m(x,t)| \leq \frac{M^m (x-t)^{m-1}}{(m-1)!} \leq \frac{M^m (b-a)^{m-1}}{(m-1)!}$$

hold for all $m \geq 1$ and arbitrary $x,t \in [a,b]$. Given these estimates, each term of the sum in Eq. (4.2) satisfies the inequality

$$\left| \lambda^m \int_a^x K_m(x,t) f(t)\,\mathrm{d}t \right| \leq \frac{|\lambda|^m M^m (b-a)^{m-1}}{(m-1)!} \|f\|_1.$$

Hence, the iterative sequence $\{\phi_n(x)\}$ of approximants converges absolutely and uniformly to the continuous solution

$$\phi(x) = f(x) + \sum_{m=1}^{\infty} \lambda^m \left(\int_a^x K_m(x,t) f(t)\,\mathrm{d}t \right)$$

of the integral equation (4.1) for every complex value of λ, as it is dominated by an absolutely convergent series. Since the order of summation and integration can now be legitimately interchanged, the form of the solution becomes

$$\phi(x) = f(x) + \lambda \int_a^x R(x,t;\lambda) f(t)\,\mathrm{d}t,$$

where we have set

$$R(x,t;\lambda) = \sum_{m=1}^{\infty} \lambda^{m-1} K_m(x,t).$$

The infinite series $R(x,t;\lambda)$ is known as the *resolvent kernel* of the integral equation (4.1), or as the *Neumann series*.

The solution $\phi(x)$ to the integral equation is unique. If there were two solutions, $\phi(x)$ and $\tilde{\phi}(x)$, then their difference $\delta(x) = \phi(x) - \tilde{\phi}(x)$ would satisfy the homogeneous equation

$$\delta(x) = \lambda \int_a^x K(x,t)\,\delta(t)\,\mathrm{d}t.$$

We show that $\delta(x) \equiv 0$ is the only solution to this integral equation. Let

$$d^2 = \int_a^b \delta^2(x)\,\mathrm{d}x, \qquad A^2(x) = \int_a^x K(x,t)\,\mathrm{d}t, \quad \text{and} \quad N^2 = \int_a^b A^2(x)\,\mathrm{d}x.$$

An application of the Cauchy–Schwarz inequality shows that

$$\delta^2(x) \le |\lambda|^2 \left(\int_a^x K^2(x,t)\,\mathrm{d}t \right) \left(\int_a^x \delta^2(t)\,\mathrm{d}t \right) \le |\lambda|^2 A^2(x)\, d^2.$$

After replacing x by t in this inequality and then integrating from a to x, we obtain

$$\int_a^x \delta^2(t)\,\mathrm{d}t \le |\lambda|^2 \left(\int_a^x A^2(t)\,\mathrm{d}t \right) d^2.$$

If we set

$$B_1(x) = \int_a^x A^2(t)\,\mathrm{d}t$$

and combine the previous two inequalities, we obtain

$$\delta^2(x) \le |\lambda|^4 d^2 A^2(x) B_1(x).$$

If we set

$$B_2(x) = \int_a^x A^2(t) B_1(t)\,\mathrm{d}t$$

and repeat this process, we obtain

$$\delta^2(x) \le |\lambda|^6 d^2 A^2(x) B_2(x).$$

In general, if we set

$$B_n(x) = \int_a^x A^2(t) B_{n-1}(t)\,\mathrm{d}t$$

and repeat this process indefinitely, we obtain

$$\delta^2(x) \le |\lambda|^{2n+2} d^2 A^2(x) B_n(x)$$

for every $n \ge 1$. Observe that $B_n(a) = 0$ for each value of n. Also,

$$B_2(x) = \int_a^x A^2(t) B_1(t)\,\mathrm{d}t = \int_a^x B_1(t) B_1'(t)\,\mathrm{d}t = \frac{1}{2!} B_1^2(x)$$

for each $x \in [a,b]$. After employing a short inductive argument, we get

$$B_n(x) = \frac{1}{n!} B_1^n(x),$$

so that

$$|B_n(x)| \le \frac{1}{n!}\,|B_1(x)|^n \le \frac{N^{2n}}{n!}$$

for every $n \ge 1$. After applying this estimate, we obtain

$$\delta^2(x) \le |\lambda|^2 d^2 A^2(x)\,\frac{(|\lambda|N)^{2n}}{n!}$$

for each $x \in [a,b]$ and every $n \ge 1$. If $n \to \infty$, we can only conclude that $\delta(x) \equiv 0$.

The discussion above serves to establish the following result:

Theorem 4.2.1 (Successive Approximation). *Let λ be a complex parameter and let $f(x)$ be a complex-valued continuous function defined on the interval $[a,b]$. Let $K(x,t)$ be a complex-valued continuous kernel defined on the triangle $T(a,b)$, with $K(x,t) \equiv 0$ if $x < t$. Then, for every complex value of λ, the unique continuous solution to the Volterra integral equation*

$$\phi(x) = f(x) + \lambda \int_a^x K(x,t)\,\phi(t)\,\mathrm{d}t$$

is given by

$$\phi(x) = f(x) + \lambda \int_a^x R(x,t;\lambda)\,f(t)\,\mathrm{d}t,$$

where $R(x,t;\lambda)$ is the unique resolvent kernel

$$R(x,t;\lambda) = \sum_{m=1}^{\infty} \lambda^{m-1} K_m(x,t).$$

One notable consequence of this theorem is that $\phi(x) \equiv 0$ if $f(x) \equiv 0$.

Another notable consequence is that Volterra kernels have no eigenvalues, since the resolvent series is an entire function of λ.

The magnitude of the error due to the use of the approximant $\phi_n(x)$ in estimating the solution $\phi(x)$ can be uniformly estimated by using the same estimate established within the proof. For each $x \in [a,b]$, we have

$$|\phi(x) - \phi_n(x)| \le |\lambda|\,M\,\|f\|_1 \sum_{m=n}^{\infty} \frac{[|\lambda|\,M\,(b-a)^m]}{m!}.$$

The sum on the right can be estimated by Lagrange's form of the remainder for the exponential series. In doing so, we obtain the uniform estimate

$$\|\phi(x) - \phi_n(x)\|_\infty \le \mathrm{e}^b\,\frac{[|\lambda|\,M\,(b-a)]^n}{n!}.$$

Thus, the magnitude of the error will be as small as desired, provided that n is large enough.

The Method of Successive Approximation firmly establishes the equivalence between solving a Volterra integral equation of the second kind and computing the resolvent kernel $R(x,t;\lambda)$ from the given kernel $K(x,t)$. The following examples demonstrate this equivalence.

Illustrative Examples

- *Example 1*: Consider the linear Volterra integral equation

$$\phi(x) = f(x) + \lambda \int_0^x xt\,\phi(t)\,\mathrm{d}t.$$

An elementary integration shows that

$$K_2(x,t) = \int_t^x xs \cdot st\,\mathrm{d}s = xt\left(\frac{x^3 - t^3}{3}\right).$$

An inductive argument shows that in general, we have

$$K_m(x,t) = \frac{xt}{(m-1)!}\left(\frac{x^3 - t^3}{3}\right)^{m-1}.$$

Thus, the resolvent kernel is

$$R(x,t;\lambda) = \sum_{m=1}^{\infty} \lambda^{m-1} K_m(x,t) = xt\,\exp\left\{\lambda\left(\frac{x^3 - t^3}{3}\right)\right\}.$$

As a consequence of the theorem, the solution to the integral equation is

$$\phi(x) = f(x) + \lambda \int_0^x xt\,\exp\left\{\lambda\left(\frac{x^3 - t^3}{3}\right)\right\} f(t)\,\mathrm{d}t.$$

In particular, if $f(x) = x$ and $\lambda = 1$, then the solution to the equation

$$\phi(x) = x + \int_0^x xt\,\phi(t)\,\mathrm{d}t$$

is

$$\phi(x) = x + \int_0^x xt\,\exp\left\{\left(\frac{x^3 - t^3}{3}\right)\right\} t\,\mathrm{d}t = x\mathrm{e}^{x^3/3}.$$

- *Example 2*: If a kernel has the simple separable form $K(x,t) = a(x)\,b(t)$, then its iterated kernels are easily computed. Indeed,

$$\begin{aligned} K_2(x,t) &= \int_t^x a(x)\,b(s)\,a(s)\,b(t)\,\mathrm{d}s \\ &= K(x,t)\int_t^x K(s,s)\,\mathrm{d}s \\ &= K(x,t)\,(L(x)-L(t)), \end{aligned}$$

where $L(s)$ is any antiderivative of $K(s,s)$. Another iteration shows that

$$K_3(x,t) = K(x,t)\,\frac{(L(x)-L(t))^2}{2!}.$$

A short inductive argument shows that in general

$$K_n(x,t) = K(x,t)\,\frac{(L(x)-L(t))^{n-1}}{(n-1)!}.$$

Consequently, the resolvent kernel is given by

$$R(x,t;\lambda) = K(x,t)\,\exp\{\lambda\,(L(x)-L(t))\}.$$

Note that all of the iterated kernels and the resolvent kernel are also separable. Thus, the solution to the integral equation

$$\phi(x) = f(x) + \lambda\int_0^x a(x)\,b(t)\,\phi(t)\,\mathrm{d}t$$

takes the form

$$\phi(x) = f(x) + \lambda\int_0^x a(x)\,b(t)\,\exp\{\lambda\,(L(x)-L(t))\}\,f(t)\,\mathrm{d}t.$$

In the previous example, $K(x,t) = xt$. By using this method, we compute the antiderivative

$$L(s) = \int K(s,s)\,\mathrm{d}s = \int s^2\,\mathrm{d}s = \frac{s^3}{3},$$

so that

$$R(x,t;\lambda) = xt\exp\left\{\lambda\left(\frac{x^3-t^3}{3}\right)\right\},$$

which is exactly what had been previously obtained.

For another example of the utility of this method, consider the simple kernel $K(x,t) = \mathrm{e}^{x-t}$. Since $K(s,s) = 1$, $L(s) = s$. Therefore, the resolvent kernel is given by

$$R(x,t;\lambda) = \mathrm{e}^{x-t}\,\mathrm{e}^{\lambda\,(x-t)} = \mathrm{e}^{(\lambda+1)\,(x-t)}.$$

- *Example 3*: Consider the linear Volterra integral equation

$$\phi(x) = f(x) + \lambda \int_0^x \frac{1}{1+x}\,\phi(t)\,\mathrm{d}t.$$

Since $K(s,s) = 1/(1+s)$, $L(s) = \ln(1+s)$. Following the prescription in the previous example, we get

$$R(x,t;\lambda) = \frac{1}{1+x}\,\mathrm{e}^{\lambda\,(L(x)-L(t))} = \frac{1}{1+x}\left(\frac{1+x}{1+t}\right)^{\lambda}.$$

Therefore, the solution to the integral equation is

$$\phi(x) = f(x) + \lambda \int_0^x \frac{1}{1+x}\left(\frac{1+x}{1+t}\right)^{\lambda} f(t)\,\mathrm{d}t.$$

In particular, if $f(x) = 1$, then a routine integration gives

$$\phi(x;\lambda) = \frac{1-\lambda\,(1+x)^{\lambda-1}}{1-\lambda},$$

if $\lambda \neq 1$. An application of l'Hôpital's rule gives $\phi(x;1) = 1+\ln x$ as $\lambda \to 1$, and this solution can be independently verified.

- *Example 4*: Consider the linear Volterra integral equation

$$\phi(x) = f(x) + \mu^2 \int_0^x (x-t)\,\phi(t)\,\mathrm{d}t.$$

An elementary integration shows that

$$K_2(x,t) = \int_t^x (x-s)(s-t)\,\mathrm{d}s = \frac{1}{3!}\,(x-t)^3.$$

An inductive argument shows that in general, we have

$$K_m(x,t) = \frac{1}{(2m-1)!}\,(x-t)^{2m-1}.$$

The resolvent kernel is

$$\begin{aligned} R(x,t;\lambda) &= \sum_{m=1}^{\infty} \mu^{2(m-1)} K_m(x,t) \\ &= \frac{1}{\mu} \sum_{m=0}^{\infty} \frac{1}{(2m+1)!} (\mu(x-t))^{2m+1} \\ &= \frac{1}{\mu} \sinh(\mu(x-t)). \end{aligned}$$

As a consequence of the theorem, the solution to the equation is

$$\phi(x) = f(x) + \mu \int_0^x \sinh(\mu(x-t)) f(t)\,\mathrm{d}t.$$

- *Example 5*: In the proof of the theorem, the initial approximant $\phi_0(x)$ was chosen to be $f(x)$. However, if the integration necessary to compute $\phi_1(x)$ is difficult, then an alternate choice for $\phi_0(x)$ may be more practicable—the process of integration may be less difficult, or the rate of convergence may be accelerated. The purpose of this example is to illustrate a nice technique by which this can be accomplished.

  Suppose that we wish to compute a sequence $\{\phi_n(x)\}$ of approximants to the solution to the Volterra integral equation

$$\phi(x) = 1 + \int_0^x \frac{1}{1-2xt+t^2} \phi(t)\,\mathrm{d}t$$

on the interval $[0, \frac{1}{2}]$. If we choose $\phi_0(x) = 1$, then a straightforward integration gives

$$\phi_1(x) = 1 + \frac{1}{\sqrt{1-x^2}} \arctan\left(\frac{x}{\sqrt{1-x^2}}\right).$$

Since computing $\phi_2(x)$ would be prohibitively difficult, we seek a better choice for $\phi_0(x)$.

Note first that $\phi(0) = 1$. If we apply the tool for differentiating an integral that was described in Sect. 4.1, then we obtain

$$\phi'(x) = \frac{1}{1-x^2} \phi(x) + \int_0^x \frac{2t}{(1-2xt+t^2)^2} \phi(t)\,\mathrm{d}t,$$

from which we obtain $\phi'(0) = 1$. If we differentiate the integral equation twice more, then we obtain $\phi''(0) = 1$ and $\phi'''(0) = 5$. Thus, the first four terms of the Maclaurin series for the solution are

$$\phi(x) = 1 + x + \frac{1}{2}x^2 + \frac{5}{6}x^3 + \cdots$$

As a zeroth approximant to the solution $\phi(x)$, we choose

$$\psi_0(x) = \phi(x) = 1 + x + \frac{1}{2}x^2 + \frac{5}{6}x^3.$$

Now let $P_7(x,t)$ denote the seventh partial sum of the kernel $1/(1-2xt+t^2)$. Then

$$\begin{aligned} P_7(x,t) = 1 &+ 2xt + (4x^2-1)t^2 + (8x^3-4x)t^3 \\ &+ (16x^4-12x^2+1)t^4 + (32x^5-32x^3+6x)t^5 \\ &+ (64x^6-80x^4+24x^2-1)t^6 + (128x^7-192x^5+80x^3-8x)t^7. \end{aligned}$$

Since $P_7(x,t)$ agrees with the kernel up to the power t^7, we can replace the kernel with $P_7(x,t)$ in our calculations to obtain

$$\begin{aligned} \psi_1(x) &= 1 + x + \frac{1}{2}x^2 + \frac{5}{6}x^3 + \frac{5}{8}x^4 + \frac{41}{60}x^5 + \frac{101}{180}x^6 + \frac{25}{42}x^7 + O\left(x^8\right) \\ \psi_2(x) &= 1 + x + \frac{1}{2}x^2 + \frac{5}{6}x^3 + \frac{5}{8}x^4 + \frac{97}{120}x^5 + \frac{27}{40}x^6 + \frac{143}{180}x^7 + O\left(x^8\right) \\ \psi_3(x) &= 1 + x + \frac{1}{2}x^2 + \frac{5}{6}x^3 + \frac{5}{8}x^4 + \frac{97}{120}x^5 + \frac{167}{240}x^6 + \frac{227}{280}x^7 + O\left(x^8\right) \\ \psi_4(x) &= 1 + x + \frac{1}{2}x^2 + \frac{5}{6}x^3 + \frac{5}{8}x^4 + \frac{97}{120}x^5 + \frac{167}{240}x^6 + \frac{1367}{1680}x^7 + O\left(x^8\right). \end{aligned}$$

Each approximant $\psi_n(x)$ is a polynomial. The fifth approximant $\psi_5(x)$ agrees with $\psi_4(x)$ up to the term containing x^7. These approximants will agree with $\phi(x)$ up to the term containing x^7, since the approximate kernel $P_7(x,t)$ agrees with the kernel that far. If the last three approximants are plotted in the same plane, their graphs are indistinguishable to the naked eye. A painstaking technical analysis would reveal that the errors involved here are quite small on the interval $[0, \frac{1}{2}]$.

- *Example 6*: Consider the Volterra integral equation

$$\phi(x) = (1+x^2)^2 + \lambda \int_0^x \frac{1+x^2}{1+t^2}\,\phi(t)\,\mathrm{d}t.$$

Note that if $\lambda = 0$, then $\phi(x;0) = (1+x^2)^2$.

If we utilize the technique in *Example 2* to compute the corresponding resolvent kernel, then the solution to the integral equation assumes the form

$$\begin{aligned} \phi(x;\lambda) &= (1+x^2)^2 + \lambda \int_0^x \frac{1+x^2}{1+t^2}\,\mathrm{e}^{\lambda(x-t)}\,(1+t^2)^2\,\mathrm{d}t \\ &= \frac{1+x^2}{\lambda^2}\left[(2+\lambda^2)\,\mathrm{e}^{\lambda x} - 2(1+\lambda x)\right]. \end{aligned}$$

At first glance, it would appear that this solution is undefined at $\lambda = 0$; however, after two applications of l'Hôpital's rule, we find that

$$\lim_{\lambda \to 0} \phi(x;\lambda) = \phi(x;0).$$

There is another way to evaluate this limit. Since the Maclaurin expansion for the solution is given by

$$\phi(x;\lambda) = (1+x^2)\left[1+\lambda x+\frac{1}{2}(2+\lambda^2)x^2+\frac{1}{6}\lambda(1+\lambda^2)x^3 + \frac{1}{24}\lambda^2(2+\lambda^2)x^4+\frac{1}{120}\lambda^3(2+\lambda^2)x^5+O(x^6)\right].$$

it again follows that $\phi(x;\lambda) \to \phi(x;0)$ as $\lambda \to 0$.

Section 4.2 Exercises

1. The Theorem of Successive Approximation was proven assuming that $f(x)$ is continuous on the interval $[a,b]$ and that $K(x,t)$ is continuous on the square $Q(a,b)$. The resolvent kernel will converge absolutely and uniformly under weakened hypotheses. Investigate the sequence $\phi_n(x)$ and the convergence of the resolvent series if neither $f(x)$ nor $K(x,t)$ is necessarily continuous, but

   (a) $\|f\|_1 < +\infty$ and $K(x,t)$ is bounded.
   (b) $\|f\|_2 < +\infty$ and $\|K\|_2 < +\infty$.

   State and prove two theorems analogous to the Theorem of Successive Approximation in these cases.
2. By using its series representation, show that the resolvent kernel satisfies the integral equation

$$R(x,t;\lambda) = K(x,t) + \lambda \int_t^x K(x,s)R(s,t;\lambda)\,\mathrm{d}s$$

   for all complex values of λ.
3. For each of the given kernels, compute the corresponding resolvent kernels by following the prescription given in *Example 2*:

   (a) $A(x,t) = \sqrt{\cos x \cos t}$
   (b) $B(x,t) = \sqrt{\sin x \sin t}$
   (c) $C(x,t) = \sqrt{\tan x \tan t}$
   (d) $D(x,t) = 1/\sqrt{xt}$
   (e) $E(x,t) = x^3 t^3$

(f) $F(x,t) = (3x+1)/(3t+1)^2$
(g) $G(x,t) = 2\sin x \cos t$
(h) $H(x,t) = \ln x / \ln t$

4. Consider the Volterra integral equation

$$\phi(x) = \frac{1}{1+x^2} + \lambda \int_0^x \frac{t}{1+x^2}\,\phi(t)\,\mathrm{d}t.$$

Follow the prescription given in *Example 2*, to show that the resolvent kernel is given by

$$R(x,t;\lambda) = \frac{t}{1+x^2}\left(\frac{1+x^2}{1+t^2}\right)^{\lambda/2}.$$

Use this representation to show that the solution is given by

$$\phi(x) = \left(1+x^2\right)^{(\lambda-2)/2}.$$

Sketch the solution for various values of λ.

5. Use the results of Problem 3 to solve the following integral equations:

(a) $\phi(x) = \cos^{1/2} x + \lambda \int_0^x \sqrt{\cos x}\sqrt{\cos t}\,\phi(t)\,\mathrm{d}t$

Answer: $\phi(x) = \sqrt{\cos x}\,\mathrm{e}^{\lambda \sin x}$.

(b) $\phi(x) = \sin^{1/2} x + \lambda \int_0^x \sqrt{\sin x}\sqrt{\sin t}\,\phi(t)\,\mathrm{d}t$

Answer: $\phi(x) = \sqrt{\sin x}\,\mathrm{e}^{-\lambda \cos x}$.

(c) $\phi(x) = \sqrt{\sin x} + \frac{1}{2}\int_0^x \sqrt{\tan x}\sqrt{\tan t}\,\phi(t)\,\mathrm{d}t$

Answer: $\phi(x) = \sqrt{\sin x} + \frac{1}{2}\sqrt{\tan x}\sqrt{\sec x}\left(\frac{1}{2} - \frac{1}{2}\cos x\right)$.

(d) $\phi(x) = x^n + \lambda \int_1^x (1/\sqrt{xt})\,\phi(t)\,\mathrm{d}t$

Answer: If $\lambda \neq n+\frac{1}{2}$, then $\phi(x) = \dfrac{(n+\frac{1}{2})x^n - \lambda x^{\lambda-1/2}}{n+\frac{1}{2}-\lambda}$.

If $\lambda = n+\frac{1}{2}$, then $\phi(x) = x^n + (n+\frac{1}{2})x^n \ln x$.

In this problem, what is the effect of changing the lower limit of integration from 1 to 0?

(e) $\phi(x) = x^3 + \lambda \int_0^x x^3 t^3\,\phi(t)\,\mathrm{d}t$

Answer: $\phi(x) = x^3 \exp\{\lambda x^7\}$.

(f) $\phi(x) = 1 + \lambda \int_0^x \frac{3x+1}{(3t+1)^2}\,\phi(t)\,dt$

Answer: If $\lambda \neq -3$, then $\phi(x) = \frac{3+\lambda\,(3x+1)^{(\lambda+3)/3}}{3+\lambda}$.
If $\lambda = -3$, then $\phi(x) = 1 - \ln(3x+1)$.

(g) $\phi(x) = \sin x + \lambda \int_0^x 2\sin x \cos t\,\phi(t)\,dt$

Answer: $\phi(x) = \sin x \exp\{\lambda\,\sin^2 x\}$.

(h) $\phi(x) = \ln x + \lambda \int_e^x (\ln x/\ln t)\,\phi(t)\,dt$

Answer: $\phi(x) = \ln x \exp\{\lambda(x-e)\}$.

6. Explain why, without loss of generality, the lower limit a of integration in an integral equation can be replaced with zero if so desired.
7. In the proof of the Theorem of Successive Approximation, we chose $\phi_0(t) = f(t)$. What happens to the sequence $\{\phi_n(x)\}$ if we choose otherwise, for instance, if we choose $\phi_0(x)$ to be a partial sum of the series representation for $\phi(x)$? Does the sequence of approximants still converge? Use *Example 5* as a guide in your investigation.
8. Consider the Volterra integral equation

$$\phi(x) = 1 + \int_0^x (x^2 - t^2)\,\phi(t)\,dt.$$

(a) Choose $\phi_0(x) = 1$. Show that the next three successive approximants to the solution of the equation are

$$\phi_1(x) = 1 + \frac{2}{3}x^3$$

$$\phi_2(x) = 1 + \frac{2}{3}x^3 + \frac{1}{12}x^6$$

$$\phi_3(x) = 1 + \frac{2}{3}x^3 + \frac{1}{12}x^6 + \frac{1}{378}x^9.$$

(b) Prove by induction that $\phi_n(x)$ is a polynomial in x^3 for every $n \geq 1$.
(c) Explain why $\phi(x)$ is a function of x^3.
(d) Substitute the power series

$$\phi(x) = \sum_{n=0}^{\infty} a_n x^{3n}$$

into the integral equation and deduce that

$$a_{n+1} = \frac{2}{(3n+1)(3n+3)}\, a_n$$

for every $n \geq 0$.

(e) With $a_0 = \phi(0) = 1$, solve this recurrence relation for a_n in terms of n.

(f) The modified Bessel function $I_\nu(t)$ of the first kind of order ν has the series representation

$$I_\nu(t) = \sum_{m=0}^{\infty} \frac{1}{2^{2m+\nu}\, m!\, \Gamma(m+\nu+1)}\, x^{2m+\nu},$$

where $\Gamma(x)$ is Euler's gamma function. Use the identity $\Gamma(m+\nu) = (\nu)_m \Gamma(\nu)$ to show that the solution to the integral equation is

$$\phi(x) = \frac{2^{1/3}\,\Gamma\left(\frac{1}{3}\right)}{3^{2/3}}\, x I_{-2/3}\left(\frac{2^{3/2} x^{3/2}}{3}\right).$$

4.3 Convolution Kernels

If the kernel $K(x,t)$ of a Volterra integral equation is a function of the difference $x-t$, i.e., if $K(x,t) = k(x-t)$ for some function k, then it is called *convolution kernel*.

If $K(x,t)$ is a convolution kernel, then the computed iterated kernels and the resolvent kernel are also convolution kernels. To see that this is true, suppose that $K(x,t) = K_1(x,t) = k_1(x-t)$ for some function k_1. Then a simple change of variable $(u = s-t)$ shows that the iterated kernel

$$\begin{aligned} K_2(x,t) &= \int_t^x k_1(x-s)\, k_1(s-t)\, \mathrm{d}s \\ &= \int_0^{x-t} k_1((x-t)-u)\, k_1(u)\, \mathrm{d}u \\ &= k_2(x-t) \end{aligned}$$

for some function k_2. An inductive argument shows that $K_m(x,t) = k_m(x-t)$ for some function k_m, so that

$$\begin{aligned} R(x,t;\lambda) &= \sum_{m=1}^{\infty} \lambda^{m-1} K_m(x,t) \\ &= \sum_{m=1}^{\infty} \lambda^{m-1} k_m(x-t) \\ &= r(x-t;\lambda) \end{aligned}$$

for some function $r(x;\lambda)$.

The Laplace transform of $r(x;\lambda)$ can be expressed in terms of the Laplace transform of $k(x)$. According to the Theorem of Successive Approximation, the solution to the equation

$$\phi(x) = f(x) + \lambda \int_0^x k(x-t)\,\phi(t)\,\mathrm{d}t \tag{4.3}$$

has the representation

$$\phi(x) = f(x) + \lambda \int_0^x r(x-t;\lambda)\,f(t)\,\mathrm{d}t. \tag{4.4}$$

Note that if $\mathcal{L}\{k(x)\}$ and $\mathcal{L}\{\phi(x)\}$ exist, then

$$\mathcal{L}\left\{\int_0^x k(x-t)\,\phi(t)\,\mathrm{d}t\right\} = \mathcal{L}\{(k*\phi)(x)\} = \mathcal{L}\{k(x)\}\,\mathcal{L}\{\phi(x)\}.$$

Applying the Laplace transform to Eq. (4.3), we get

$$\mathcal{L}\{\phi(x)\} = \mathcal{L}\{f(x)\} + \lambda\,\mathcal{L}\{k(x)\}\,\mathcal{L}\{\phi(x)\}, \tag{4.5}$$

and applying the Laplace transform to Eq. (4.4), we get

$$\mathcal{L}\{\phi(x)\} = \mathcal{L}\{f(x)\} + \lambda\,\mathcal{L}\{r(x;\lambda)\}\,\mathcal{L}\{f(x)\}. \tag{4.6}$$

By eliminating $\mathcal{L}\{\phi(x)\}$ and $\mathcal{L}\{f(x)\}$ from Eqs. (4.5) and (4.6), we obtain

$$\mathcal{L}\{r(x;\lambda)\} = \frac{\mathcal{L}\{k(x)\}}{1-\lambda\,\mathcal{L}\{k(x)\}}, \tag{4.7}$$

from which $r(x;\lambda)$ can be determined.

The Laplace transform can be used to solve Volterra integral equations with convolution kernels of the form (4.3). By solving Eq. (4.5) for $\mathcal{L}\{\phi(x)\}$, we obtain

$$\mathcal{L}\{\phi(x)\} = \frac{\mathcal{L}\{f(x)\}}{1-\lambda\,\mathcal{L}\{k(x)\}}.$$

The solution to the integral equation is the inverse Laplace transform of the right-hand side of this equation.

Illustrative Examples

- *Example 1*: Consider the integral equation

$$\phi(x) = 2 + \lambda^4 \int_0^x \frac{(x-t)^3}{3!}\,\phi(t)\,\mathrm{d}t.$$

After applying the Laplace transform to this equation and then solving for $\mathcal{L}\{\phi(x)\}$, we obtain

$$\begin{aligned}\mathcal{L}\{\phi(x)\} &= \frac{2s^3}{s^4-\lambda^4} \\ &= \frac{1}{2}\,\frac{1}{s-\lambda} + \frac{1}{2}\,\frac{1}{s+\lambda} + \frac{s}{s^2+\lambda^2}.\end{aligned}$$

After inverting each of these terms, we obtain

$$\phi(x) = \tfrac{1}{2}\,\mathrm{e}^{\lambda t} + \tfrac{1}{2}\,\mathrm{e}^{-\lambda t} + \cos(\lambda t) = \cosh(\lambda t) + \cos(\lambda t),$$

which is the unique solution to the integral equation for all values of λ.

- *Example 2*: For real values of λ, consider the integral equation

$$\phi(x) = f(x) + \lambda \int_0^x \sin(x-t)\,\phi(t)\,\mathrm{d}t.$$

Since $\mathcal{L}\{\sin(x)\} = 1/(s^2+1)$, an application of formula (4.7) above shows that $\mathcal{L}\{r(x;\lambda)\} = 1/(s^2+1-\lambda)$. The solution to this integral equation is

$$\phi(x) = f(x) + \lambda \int_0^x r(x-t;\lambda)\,f(t)\,\mathrm{d}t,$$

where

$$r(x;\lambda) = \begin{cases} \dfrac{\sin(\sqrt{1-\lambda}\,x)}{\sqrt{1-\lambda}}, & \text{if } -\infty < \lambda < 1; \\ x, & \text{if } \lambda = 1; \\ \dfrac{\sinh(\sqrt{\lambda-1}\,x)}{\sqrt{\lambda-1}}, & \text{if } 1 < \lambda < +\infty. \end{cases}$$

Note that the structural form of the solution changes near $\lambda = 1$, but that

$$\lim_{\lambda \to 1^-} r(x;\lambda) = r(x;1) = \lim_{\lambda \to 1^+} r(x;\lambda).$$

- *Example 3*: Consider the integral equation

$$\phi(x) = \sin(2x) + \mu^2 \int_0^x (x-t)\,\phi(t)\,\mathrm{d}t.$$

After applying the Laplace transform to this equation and then solving for $\mathcal{L}\{\phi(x)\}$, we obtain

$$\begin{aligned} \mathcal{L}\{\phi(x)\} &= \frac{2s^2}{(s^2-\mu^2)(s^2+4)} \\ &= \frac{\mu}{\mu^2+4}\,\frac{1}{s-\mu} - \frac{\mu}{\mu^2+4}\,\frac{1}{s+\mu} + \frac{4}{\mu^2+4}\,\frac{2}{s^2+4}. \end{aligned}$$

After inverting each of these terms, we obtain

$$\begin{aligned} \phi(x;\mu) &= \frac{\mu}{\mu^2+4}\,\mathrm{e}^{\mu x} - \frac{\mu}{\mu^2+4}\,\mathrm{e}^{-\mu x} + \frac{4}{\mu^2+4}\,\sin(2x) \\ &= \frac{2\mu}{\mu^2+4}\,\sinh(\mu x) + \frac{4}{\mu^2+4}\,\sin(2x), \end{aligned}$$

which is the unique solution to the integral equation above provided that $\mu \neq \pm 2\mathrm{i}$. If $\mu = \pm 2\mathrm{i}$ or $\mu^2 = -4$, then

$$\lim_{\mu \to \pm 2\mathrm{i}} \phi(x;\mu) = x\cos(2x) + \frac{1}{2}\,\sin(2x) = \phi(x;\pm 2\mathrm{i}).$$

The last equality is justified here, since it can be independently verified as the unique solution to the equation when $\mu = \pm 2\mathrm{i}$. Thus, the solution to this equation is a continuous function of the parameter μ.

This equation can be solved by another method. Since $\mathcal{L}\{k(x)\} = 1/s^2$, an application of formula (4.7) gives $\mathcal{L}\{r(x;\mu^2)\} = 1/(s^2-\mu^2)$, so that $r(x;\mu^2) = \sinh(\mu(x))/\mu$. Hence, the solution to the integral equation takes the form

$$\phi(x;\mu) = \sin(2x) + \mu \int_0^x \sinh(\mu(x-t))\sin(2t)\,\mathrm{d}t,$$

which is equivalent to the solution obtained above.

Section 4.3 Exercises

1. Let $F(x,t)$ be a complex-valued continuous function that is defined on the square $Q(a,b)$ and is continuously differentiable with respect to x. Justify the following limiting processes:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}\left(\int_a^x F(x,t)\,\mathrm{d}t\right) &= \lim_{h\to 0}\left[\frac{1}{h}\left(\int_a^{x+h} F(x+h,t)\,\mathrm{d}t - \int_a^x F(x,t)\,\mathrm{d}t\right)\right] \\ &= \lim_{h\to 0}\left[\frac{1}{h}\int_x^{x+h} F(x+h,t)\,\mathrm{d}t\right] \\ &\quad + \lim_{h\to 0}\left[\int_a^x \left(\frac{F(x+h,t)-F(x,t)}{h}\right)\right] \\ &= F(x,x) + \int_a^x \frac{\partial F}{\partial x}(x,t)\,\mathrm{d}t.\end{aligned}$$

2. Consider the Volterra integral equation

$$\phi(x) = a + bx + \lambda \int_0^x [c + d(x-t)]\,\phi(t)\,\mathrm{d}t,$$

where a, b, c, and d are arbitrary constants. Categorize all possible solutions to this integral equation.

Hint: Show that

$$\mathcal{L}\{\phi(x)\} = \frac{as+b}{s^2 - \lambda c s - \lambda d}.$$

What are all possible partial fraction decompositions of this transform?

3. Solve each of the following integral equations:

   (a) $\phi(x) = 3 + 4x + \int_0^x [2 - (x-t)]\,\phi(t)\,\mathrm{d}t$
   Answer: $\phi(x) = (3+7x)\,\mathrm{e}^x$.

   (b) $\phi(x) = 3 + 4x + \int_0^x [2 - 2(x-t)]\,\phi(t)\,\mathrm{d}t$
   Answer: $\phi(x) = \mathrm{e}^x\,(3\cos x + 7\sin x)$.

4. Consider the Volterra integral equation

$$\phi(x) = a + bx + cx^2 + \lambda \int_0^x [d + e(x-t) + f(x-t)^2]\,\phi(t)\,\mathrm{d}t,$$

   where a, b, c, d, e, and f are arbitrary constants. Categorize all possible solutions to this integral equation. Generalize your result to all integral equations with kernels of the form

$$K(x,t) = \sum_{k=0}^{n} \frac{a_k}{k!}\,(x-t)^k.$$

5. Solve each of the following integral equations:

   (a) $\phi(x) = 1 - x - x^2 + \int_0^x \left(2 - (x-t) + (x-t)^2\right)\,\phi(t)\,\mathrm{d}t$
   Answer: $\phi(x) = \cos x + \sin x$.

   (b) $\phi(x) = 1 + x^2 + \int_0^x \left(1 - 2(x-t) + (x-t)^2\right)\,\phi(t)\,\mathrm{d}t$
   Answer: $\phi(x) = \mathrm{e}^x$.

   (c) $\phi(x) = 1 - x + \int_0^x \left(3 - 4(x-t) + (x-t)^2\right)\,\phi(t)\,\mathrm{d}t$
   Answer: $\phi(x) = \mathrm{e}^x(\sin x + \cos x)$.

6. Solve each of the following integral equations:

   (a) $\phi(x) = \sin x + \int_0^x \phi(t)\,\mathrm{d}t$

   Hint: Since the kernel $K(x,t) \equiv 1$, it is trivially a convolution kernel.
   Answer: $\phi(x) = \frac{1}{2}\,(\mathrm{e}^t + \sin t - \cos t)$.

   (b) $\phi(x) = \mathrm{e}^{ax} + \lambda \int_0^x \mathrm{e}^{b(x-t)}\,\phi(t)\,\mathrm{d}t$
   Answer: If $\lambda \neq a - b$, then

$$\phi(x;\lambda) = \frac{(a-b)\,\mathrm{e}^{ax} - \lambda\,\mathrm{e}^{(b+\lambda)x}}{a-b-\lambda}.$$

   If $\lambda = a - b$, then $\phi(x; a-b) = \mathrm{e}^{ax}\,[1 + (a-b)x]$.

(c) $\phi(x) = 1 + \frac{7}{4}\int_0^x \sin(4(x-t))\,\phi(t)\,\mathrm{d}t$
Answer: $\phi(x) = \frac{1}{9}(16 - 7\cos(3x))$.
(d) $\phi(x) = \cosh x + 3\int_0^x \sinh(x-t)\,\phi(t)\,\mathrm{d}t$
Answer: $\phi(x) = \cosh(2x)$.
(e) $\phi(x) = \mathrm{e}^x + \mu^2 \int_0^x (x-t)\,\mathrm{e}^{x-t}\,\phi(t)\,\mathrm{d}t$
Answer: $\phi(x) = \mathrm{e}^x \cosh(\mu x)$.
(f) $\phi(x) = \sin x + \sqrt{5}\int_0^x \cos(x-t)\,\phi(t)\,\mathrm{d}t$
Answer: $\phi(x) = 2\,\mathrm{e}^{x\sqrt{5}/2}\sinh(x/2)$.
(g) $\phi(x) = x + 216\int_0^x (x-t)^3\,\phi(t)\,\mathrm{d}t$
Answer: $\phi(x) = \frac{1}{12}[\sinh(6x) + \sin(6x)]$.
(h) Let n be a positive integer. Solve the integral equations

$$\phi(x) = \tfrac{1}{k!}\,x^k + \tfrac{1}{n!}\int_0^x (x-t)^n\,\phi(t)\,\mathrm{d}t.$$

Hint: Solve these equations for $n = 1, 2$, and 3 and then generalize your results.

4.4 Equations of the First Kind

The *Volterra integral equation of the first kind* is an equation of the form

$$\int_a^x K(x,t)\,\phi(t)\,\mathrm{d}t = f(x). \tag{4.8}$$

In order to guarantee the existence of a unique continuous solution $\phi(x)$ to this integral equation on the interval $[a,b]$, we require that $f(x)$ be continuous on that interval and that $K(x,t)$ be continuous on the triangle $T(a,b)$. A simple example illustrates the continuity requirement.

Consider the Volterra integral equation

$$\int_0^x t^{x-t}\,\phi(t)\,\mathrm{d}t = f(x).$$

The kernel $K(x,t) = t^{x-t}$ is continuous on the triangle $T(0,b)$. If $f(x) \equiv 0$ on the interval $[0,b]$, then $\phi(x) \equiv 0$ is clearly the unique solution to the equation. However, if $f(x) = x^{x-1}$, then $f(x)$ is continuous on the interval $(0,b]$, but $f(0)$ is undefined since $f(x) = O(1/x)$ as $x \to 0$. Since we have

$$\int_0^x t^{x-t}\, t^{t-1}\, \mathrm{d}t = \left. \frac{t^x}{x} \right|_{t=0}^{x} = x^{x-1},$$

the solution $\phi(x) = x^{x-1}$ to the given integral equation is continuous on $(0,b]$.

Two of the standard analytical approaches to the resolution of a Volterra integral equation of the first kind with a general kernel involve its reduction to an integral equation of the second kind.

The first approach depends upon the the differentiation tool supplied in Sect. 4.1 that can be directly applied to obtain

$$K(x,x)\,\phi(x) + \int_a^x \frac{\partial K(x,t)}{\partial x}\,\phi(t)\,\mathrm{d}t = f'(x). \tag{4.9}$$

If $K(x,x)$ does not vanish on the interval $[a,b]$, then this equation can always be reformulated as an equation of the second kind in the form

$$\phi(x) = \frac{f'(x)}{K(x,x)} - \int_a^x \left[\frac{\partial K(x,t)}{\partial x} / K(x,x) \right] \phi(t)\,\mathrm{d}t.$$

The Theorem of Successive Approximation in Sect. 4.2 assures that this equation has a unique continuous solution $\phi(x)$ on the interval $[a,b]$, provided that $f'(x)$ is continuous on the interval $[a,b]$ and that $\partial K(x,t)/\partial x$ is continuous on the triangle $T(a,b)$. Since Eq. (4.9) was obtained from Eq. (4.8) by the process of differentiation, Eq. (4.8) can be obtained from Eq. (4.9) by an integration and the additional fact that $f(a) = 0$.

The second approach depends upon the elementary technique commonly known as integration by parts. If we set

$$\Phi(x) = \int_a^x \phi(t)\,\mathrm{d}t,$$

then $\Phi(a) = 0$ and $\Phi'(x) = \phi(x)$. After applying this procedure to the integral in Eq. (4.8), we obtain

$$K(x,x)\,\Phi(x) = f(x) + \int_a^x \frac{\partial K(x,t)}{\partial t}\,\Phi(t)\,\mathrm{d}t.$$

If $K(x,x)$ does not vanish on the interval $[a,b]$, then this equation can always be reformulated as

$$\Phi(x) = \frac{f(x)}{K(x,x)} + \int_a^x \left[\frac{\partial K(x,t)}{\partial t} / K(x,x) \right] \Phi(t)\,\mathrm{d}t.$$

The Theorem of Successive Approximation assures us that this equation has a unique continuous solution $\Phi(x)$ on the interval $[a,b]$ that can be differentiated to obtain $\phi(x)$, provided that $\partial K(x,t)/\partial t$ is continuous on the triangle $T(a,b)$.

The success of these two approaches depends upon the assumption that $K(x,x)$ does not vanish on the interval $[a,b]$. For example, if $K(x,t) = \cos(x-t)$, then $K(x,x) \equiv 1$.

However, if $K(a,a) = 0$ or if $K(x_i,x_i) = 0$ at isolated points within the interval $[a,b]$, then these techniques cannot be employed.

Note: A Volterra integral equation of the general form

$$A(x)\,\phi(x) = f(x) + \int_a^x K(x,t)\,\phi(t)\,\mathrm{d}t,$$

where $A(x)$ vanishes at isolated points within the interval $[a,b]$, was referred to by Picard as a *Volterra integral equation of the third kind*. The analysis of these equations is especially delicate and will not be considered in this text.

Another possibility is that $K(x,x) \equiv 0$ on the interval $[a,b]$. For example, if $K(x,t) = \sin(x-t)$, then $K(x,x) \equiv 0$. In this case, Eq. (4.9) is still an equation of the first kind, and it takes the simplified form

$$\int_a^x \frac{\partial K(x,t)}{\partial x}\,\phi(t)\,\mathrm{d}t = f'(x).$$

If we assume that $f'(a) = 0$, $f'(x)$ is continuously differentiable on the interval $[a,b]$, and that the partial derivative $\partial K(x,t)/\partial x$ is continuously differentiable on the triangle $T(a,b)$, then this equation will have a unique continuous solution on the interval $[a,b]$. Obviously, the process of differentiation can be continued indefinitely, if we assume that $f^{(k)}(a) = 0$ for suitable values of k and that the kernel $K(x,t)$ and the free term $f(x)$ are sufficiently differentiable. Consider the Volterra integral equation

$$\int_0^x \frac{(x-t)^{n-1}}{(n-1)!}\,\phi(t)\,\mathrm{d}t = f(x).$$

If $f^{(k)}(0) = 0$ for $k = 0,\ldots,n-1$ and $f(x)$ is n-times continuously differentiable, then $\phi(x) = f^{(n)}(x)$.

Illustrative Examples

- *Example 1*: Consider the integral equation

$$\int_0^x \sin(x-t)\,\phi(t)\,\mathrm{d}t = \frac{1}{2}x\sin x,$$

where $K(x,t) = \sin(x-t)$. If the Laplace transform method that was described in the previous section is applied to this equation, then we obtain

$$\frac{1}{s^2+1}\,\mathcal{L}\{\phi(x)\} = \frac{s}{(s^2+1)^2}$$

or

$$\mathcal{L}\{\phi(x)\} = \frac{s}{s^2+1}$$

from which we conclude that $\phi(x) = \cos x$.

On the other hand, if we differentiate the integral equation by using the tool in Sect. 4.1, then we obtain the following equation:

$$\int_0^x \cos(x-t)\,\phi(t)\,\mathrm{d}t = \frac{1}{2}\,x\cos x + \frac{1}{2}\,\sin x,$$

which is again of the first kind, since $K(x,x) \equiv 0$ and $K_x(x,t) = \cos(x-t)$. Another differentiation leads us to the equation

$$\phi(x) - \int_0^x \sin(x-t)\,\phi(t)\,\mathrm{d}t = \cos x - \frac{1}{2}\,x\sin x,$$

which is an equation of the second kind, since $K_{xx}(x,t) = -\sin(x-t)$ and $K_x(x,x) = 1$. By adding this equation to the original equation, we obtain $\phi(x) = \cos x$. Thus, the equation was solved by using differentiation and elimination, rather than any direct technique.

- *Example 2*: Let $J_0(x)$ denote the Bessel function of the first kind of order zero. The solution to the Volterra integral equation

$$\int_0^x J_0(x-t)\,\phi(t)\,\mathrm{d}t = \sin x$$

is $\phi(x) = J_0(x)$, since $\mathcal{L}\{J_0(x)\} = 1/\sqrt{s^2+1}$ and

$$\mathcal{L}\left\{\int_0^x J_0(x-t)\,J_0(t)\,\mathrm{d}t\right\} = \frac{1}{s^2+1} = \mathcal{L}\{\sin x\}.$$

This interesting formula can be interpreted as a combinatorial identity. On the one hand, if we substitute the series representation

$$J_0(x) = \sum_0^\infty (-1)^n\,\frac{1}{2^{2n}(n!)^2}\,x^{2n}$$

into the convolution integral, then we obtain after some simplification

$$\int_0^x J_0(x-t)\,J_0(t)\,\mathrm{d}t = \sum_0^\infty (-1)^m A_m \,\frac{1}{(2m+1)!}\,x^{2m+1},$$

where

$$A_m = \frac{1}{2^{2m}} \sum_{k=0}^{m} \binom{2k}{k} \binom{2m-2k}{m-k}.$$

On the other hand, we also have

$$\sin x = \sum_{m=0}^{\infty} (-1)^m \,\frac{1}{(2m+1)!}\,x^{2m+1}.$$

After equating the coefficients in these two representations, we obtain the binomial coefficient identity

$$\sum_{k=0}^{m} \binom{2k}{k} \binom{2m-2k}{m-k} = 2^{2m},$$

which is valid for all $m \geq 0$. This identity is not new. Indeed, it can be directly obtained by squaring the series

$$\frac{1}{\sqrt{1-x}} = \sum_{n=0}^{\infty} \frac{1}{2^{2n}} \binom{2n}{n} x^n.$$

A general family of combinatorial identities can be established in a similar manner. If we set

$$H_k(x) = \frac{\sqrt{\pi}}{\Gamma(k)}\,\frac{1}{2^{k-1/2}}\,t^{k-1/2} J_{k-1/2}(x),$$

then $\mathcal{L}\{H_k(x)\} = 1/(s^2+1)^k$. On the one hand, if p is any nonnegative integer, then we obtain after some simplification

$$\begin{aligned}
\mathcal{L}\{(H_{1/2} * H_{p+1/2})(x)\} &= \mathcal{L}\{H_{1/2}(x)\}\,\mathcal{L}\{H_{p+1/2}(x)\} \\
&= \frac{1}{(s^2+1)^{p+1}} \\
&= \mathcal{L}\left\{ \sum_{m=0}^{\infty} (-1)^m B_{mp}\,\frac{1}{(2m+2p+1)!}\,x^{2m+2p+1} \right\},
\end{aligned}$$

where

$$B_{mp} = \binom{m+p}{p}.$$

On the other hand, since

$$J_p(x) = \left(\frac{x}{2}\right)^p \sum_{j=0}^{\infty} (-1)^j \frac{1}{j!(j+p)!} \left(\frac{x}{2}\right)^{2j},$$

we obtain after a direct computation

$$\begin{aligned} H_{1/2} * H_{p+1/2}(x) &= \int_0^x H_{1/2}(x-t) H_{p+1/2}(t)\,\mathrm{d}t \\ &= \sum_{m=0}^{\infty} (-1)^m C_{mp} \frac{1}{(2m+2p+1)!} x^{2m+2p+1}, \end{aligned}$$

where

$$C_{mp} = \frac{1}{\binom{2p}{p}} \left[\sum_{k=0}^{m} \frac{1}{2^{2m+p}} \binom{2m-2k}{m-k} \binom{2k+2p}{k\ \ p\ \ k+p} \right].$$

Since $B_{mp} = C_{mp}$ for all nonnegative integers m and p, we obtain as a consequence the binomial/trinomial coefficient identities

$$\sum_{k=0}^{m} \binom{2m-2k}{m-k} \binom{2k+2p}{k\ \ p\ \ k+p} = \binom{m+p}{p} \binom{2p}{p} 2^{2m+p}.$$

- *Example 3*: Consider the Volterra integral equation

$$\int_0^x \mathrm{e}^{x-t} \phi(t)\,\mathrm{d}t = \sin x.$$

Since e^{x-t} is a convolution kernel, we can apply the Laplace transform to this equation. After some simplification, we find that

$$\mathcal{L}\{\phi(x)\} = \frac{s-1}{s^2+1} = \mathcal{L}\{\cos x - \sin x\},$$

from which we conclude that $\phi(x) = \cos x - \sin x$ is the unique solution to the equation.

On the other hand, the integral equation

$$\int_0^x \mathrm{e}^{x-t} \phi(t)\,\mathrm{d}t = \cos x$$

does not have a continuous solution on an interval of the form $[0,b]$ for any b, since $\cos 0 \neq 0$. However, we can still apply the Laplace transform to this equation to obtain

$$\mathcal{L}\{\phi(x)\} = 1 - \frac{s}{s^2+1} - \frac{1}{s^2+1} = \mathcal{L}\{\delta(x) - \cos x - \sin x\},$$

from which we conclude that $\phi(x) = \delta(x) - \cos x - \sin x$ is a solution to the equation, where $\delta(x)$ is the Dirac δ-function.

- *Example 4*: Consider the Volterra integral equation

$$\int_0^x \left(x^2 - t^2\right) \phi(t)\,\mathrm{d}t = f(x).$$

The requirement that $f(0) = 0$ is a necessary but not sufficient for this equation to have a continuous solution on an interval of the form $[0,b]$.

If we set $\phi(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots$, then

$$\int_0^x \left(x^2 - t^2\right) \phi(t)\,\mathrm{d}t = \frac{2}{3} a_0 x^3 + \frac{1}{2} a_1 x^4 + \frac{2}{15} a_2 x^5 + \frac{1}{12} a_3 x^6 + \cdots.$$

Thus, it is necessary that $f(0) = f'(0) = f''(0) = 0$ for this equation to have a unique continuous solution. If we differentiate the integral equation

$$\int_0^x \left(x^2 - t^2\right) \phi(t)\,\mathrm{d}t = \sin\left(x^3\right)$$

with respect to x, then we obtain

$$\int_0^x 2x\,\phi(t)\,\mathrm{d}t = 3x^2 \cos\left(x^3\right).$$

After canceling x, another differentiation yields

$$\phi(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{3}{2} x \cos\left(x^3\right)\right] = \frac{3}{2}\left[\cos\left(x^3\right) - 3x^3 \sin\left(x^3\right)\right],$$

which is the unique continuous solution to the integral equation.

Section 4.4 Exercises

1. Suppose that the kernel $K(x,t) = a(x)\,b(t)$ is continuous on the triangle $T(a,b)$ and that $a(x)$ does not vanish on the interval $[a,b]$. Give a set of additional necessary conditions for the existence of a unique continuous solution to the Volterra integral equation

$$\int_0^x K(x,t)\,\phi(t)\,\mathrm{d}t = f(x).$$

   What is the general form of the solution?
2. Determine the unique continuous solution to the Volterra integral equation

$$\int_0^x \sqrt{\cos x \cos t}\,\phi(t)\,\mathrm{d}t = \tan x$$

   on the interval $[0,\pi/4]$. Can the interval of validity for this solution be extended indefinitely?
   Answer: $\phi(x) = \sec x\left(1 + \frac{3}{2}\tan^2 x\right)$.
3. Determine the continuous solution to the Volterra integral equation

$$\int_1^x \frac{1}{\sqrt{xt}}\,\phi(t)\,\mathrm{d}t = f(x),$$

   provided that $f(1) = 0$ and that $f(x)$ is continuously differentiable on the interval $[1,b]$.
   Answer: $\phi(x) = x f'(x) + \frac{1}{2} f(x)$.
4. Show that if $\phi(t) = c_0 + c_1 t + c_2 t^2 + \cdots$, then

$$\int_0^x \sin(2(x-t))\,\phi(t)\,\mathrm{d}t = c_0 x^2 + \tfrac{1}{3} c_1 x^3 + \cdots.$$

   Solve the Volterra integral equations

$$\int_0^x \sin(2(x-t))\,\phi(t)\,\mathrm{d}t = \sin^n x$$

   if $n = 2$, 3, and 4. Is there a solution if $n = 1$?
   Answer: If $n = 2$, then $\phi(x) = 1$.
   If $n = 3$, then $\phi(x) = \frac{9}{8}\sin x + \frac{5}{8}\sin(5x)$.
   If $n = 4$, then $\phi(x) = \frac{3}{4}(1 - \cos(4x))$.

4.5 Numerical Methods

Given the difficulties that arise in determining exact solutions to equations of the second kind as exhibited in the previous sections of this chapter, for instance, the computation of the resolvent kernel, the use of numerical methods assumes critical importance in the production of approximate solutions to equations of this type. It is not the intent of this text to provide a complete treatment of the convergence properties, the numerical stability, or a painstaking analysis of the errors inherent in these approximations. Instead, we refer the reader to more advanced texts or to the literature for a complete discussion of these issues.

In this section, we provide a condensed outline of the general procedures involved in the production of a sequence of approximate solutions to a Volterra integral equation of the second kind via quadrature methods. We mention some of the complications that naturally arise with the choice of weights, and the techniques required to assess the accuracy of the approximations. The methods here are considerably more complicated than the methods employed in Sect. 2.6 for the production of approximate solutions to linear Fredholm integral equations of the second kind.

As usual, we assume that $f(x)$ is continuous on the interval $[0,1]$ and that $K(x,t)$ is continuous on the triangle $T(0,1)$. If an approximate solution to the Volterra integral equation

$$\phi(x) = f(x) + \int_0^x K(x,t)\,\phi(t)\,\mathrm{d}t$$

is required on the interval $[0,1]$, then any standard quadrature method of the Newton–Cotes type that is used to approximate the integral requires a predetermined set $\{x_0, x_1, \ldots, x_n\}$ of $n+1$ equally spaced *nodes*, where $x_i = i/n$. For each x_i, $i = 0, 1, \ldots, n$, we have

$$\phi(x_i) = f(x_i) + \int_0^{x_i} K(x_i,t)\,\phi(t)\,\mathrm{d}t.$$

Clearly, $\phi(x_0) = f(x_0) = f(0)$. For each $i = 1, \ldots, n$, each integral here can be approximated by a finite sum on the interval $[0, x_i]$, so that

$$\phi(x_i) = f(x_i) + \sum_{j=0}^{i} w_{ij}\,K(x_i,t_j)\,\phi(t_j) + E(\Delta x, x_i),$$

where $t_j = j/n$, the weights w_{ij} are (usually) positive, and $E(\Delta x, x_i)$ is an error term attached to the choice of quadrature method, with $\Delta x = x_i - x_{i-1} = 1/n$. For example, if the trapezoid rule is employed, then the weights would be $w_{i0} = w_{ii} = \Delta x/2 = 1/2n$ and $w_{ij} = \Delta x = 1/n$ for $j = 1, \ldots, i-1$. Also, if n is sufficiently large, then $E(\Delta x, x_i)$ is sufficiently small.

Initially, we set $y_0 = \phi(x_0)$ as a *starting value*. Under appropriately crafted conditions, the error term $E(\Delta x, x_i)$ is relatively small for all $i = 1, \ldots, n$. If we discard it, then we arrive at the linear system

$$y_i = f(x_i) + \sum_{j=0}^{i} w_{ij} K(x_i, t_j) y_j.$$

When expanded, the equations in this system take the form

$$\begin{aligned} y_1 &= f(x_1) + w_{10} K(x_1, t_0) y_0 + w_{11} K(x_1, t_1) y_1, \\ y_2 &= f(x_2) + w_{20} K(x_2, t_0) y_0 + w_{21} K(x_2, t_1) y_1 + w_{22} K(x_2, t_2) y_2, \end{aligned}$$

and so forth. In these equations, $x_i = t_i = i/n$.

After a careful inspection of these equations, several observations can be made:

- The weights w_{ij} vary from one equation to the next.
- As n increases, the accuracy of the approximation increases.
- Each y_i can be computed recursively in a step-by-step fashion, and it depends upon y_0.

If a higher-order quadrature rule, such as Simpson's rule or Bode's rule, is adopted, then additional starting values must be specified before these equations can be solved recursively. For example, if Simpson's rule is employed, then the interval $[0, x_i]$ must necessarily be partitioned into an even number of subintervals, and y_1 must be specified before y_2 can be computed in terms of y_0 and y_1. If the interval $[0, x_i]$ is partitioned into an odd number of subintervals, then a hybrid of Simpson's rule with another quadrature rule must be constructed with weights adjusted accordingly.

After each y_i is determined, a continuous interpolating function $y_n(x)$ can be constructed on the interval $[0, 1]$ that passes through the points (x_i, y_i) for all $i = 0, 1, \ldots, n$. One way of accomplishing this task would be to employ the classical Lagrange interpolation formula

$$y_n(x) = \sum_{i=1}^{n} y_i \left(\prod_{\substack{1 \le k \le n \\ k \ne i}} \frac{x - x_k}{x_i - x_k} \right).$$

Some natural questions arise: For each n, how well does $y_n(x)$ approximate $\phi(x)$ on the interval $[0, 1]$? Does $y_n(x)$ converge uniformly to $\phi(x)$ on the interval $[0, 1]$ as $n \to +\infty$? In order to answer these and other relevant questions, it must be possible beforehand to establish the fact that the maximum *discretization error*

$$\max_{0 \le i \le n} |y_i - \phi(x_i)|$$

tends to 0 as $n \to +\infty$. This can be done, provided that the error $E(\Delta x, x_n)$ associated with the choice of quadrature method decreases to 0 rapidly enough and the weights $\{|w_{ij}|\}$ are bounded independently of n. In addition, it must be shown that any starting errors associated with a higher-order quadrature method must tend to 0 as $n \to +\infty$.

The several types of errors associated with the use of an approximative method can vary smoothly, or they can behave rather unpredictably. Thus, even if convergence of the approximants is assured, an analysis of the rate of growth of the various errors and accumulated errors may be required in order to have confidence in the accuracy of the computed approximations and the effectiveness of the choice of quadrature rule.

Furthermore, *numerical instability* can occur, depending upon the rate of convergence of the approximating sequence and the rate of error propagation. A thorough analysis of the numerical stability of an applied quadrature method can provide reassurance and insight, since it characterizes the essential features of the stable methods as well as the unstable ones, thereby furnishing valuable guidance to the user.

Finally, other useful methods, such as the *block-by-block* methods which generalize the well-known *Runge–Kutta* method, are readily available in the literature.

Chapter 5
Differential and Integrodifferential Equations

There are strong connections between the theory of integral equations and the theory of differential equations. Although there are many ways to illustrate, analyze, and interpret these connections, we can only discuss a few of them in this chapter.

In Sect. 5.1, we present several tools of the trade that are indispensible for the comprehension of the material in this chapter.

In Sect. 5.2, we demonstrate how Fredholm integral equations can be converted to boundary value problems for ordinary differential equations, and *vice versa*. Interconversion may yield a great advantage. For example, the eigenfunctions of a kernel may be difficult to determine, whereas these same eigenfunctions may be easy to determine if they are the solutions to a converted boundary value problem. Also, the boundary conditions that accompany an ordinary differential equation are imposed separately, but they can be directly incorporated the converted integral equation.

In Sect. 5.3, we demonstrate a few simple and convenient methods by which Volterra integral equations can be converted to initial value problems, and *vice versa*, and note the advantages of conversion.

In Sect. 5.4, we consider integrodifferential equations, which are considered to be hybrids of differential and integral equations, since they involve the integral and the derivative of the unknown function in possibly many ways. Since the variety of combinations is virtually infinite, we must restrict our consideration here to a few interesting and solvable examples of linear integrodifferential equations.

5.1 Tools of the Trade

In this chapter, the reader should be familiar with the following topics:

- *Linear ordinary differential equations*: A *linear ordinary differential equation of order n* assumes the form

S.M. Zemyan, *The Classical Theory of Integral Equations: A Concise Treatment*,
DOI 10.1007/978-0-8176-8349-8_5, © Springer Science+Business Media, LLC 2012

$$a_0(x)y^{(n)}(x) + a_1(x)y^{(n-1)}(x) + \cdots + a_{n-1}y'(x) + a_n(x)y(x) = g(x),$$

where $g(x)$ and the coefficient functions $a_i(x)$ are always assumed to be continuous on some interval $I = (a,b)$.

If $a_0(x_0) \neq 0$, then x_0 is called an *ordinary point*; if $a_0(x_0) = 0$, then x_0 is called a *singular point*. Often, when x_0 is an ordinary point, the differential equation is divided by $a_0(x)$; equivalently, we can assume that $a_0(x) \equiv 1$. For example, linear differential equations of the second order with an ordinary point at $x_0 = 0$ have the standard form

$$y''(x) + a_1(x)y'(x) + a_2(x)y(x) = 0.$$

Singular points may be either *regular* or *irregular*. A common example of a second-order linear differential equation with a regular singular point at $x_0 = 0$ is the Euler equation given by

$$x^2y''(x) + \alpha xy'(x) + \beta y(x) = 0.$$

Another well-known example is the Bessel equation of order ν given by

$$x^2y''(x) + xy'(x) + \left(x^2 - \nu^2\right)y(x) = 0.$$

Since differential equations with irregular singular points are quite difficult to solve, we do not consider them here.

Without loss of generality, we assume that $x_0 = 0$ in this chapter.

If $g(x)$ vanishes identically on I, then the equation is called *homogeneous*; if $g(x)$ does not vanish identically on I, then the equation is called *nonhomogeneous*. Every homogeneous linear ordinary differential equation of order n has exactly n linearly independent solutions $\{y_1(x), \ldots, y_n(x)\}$, and the most general solution $y_c(x)$ (the *complementary solution*) can be written as a linear combination of them. If the equation in question is nonhomogeneous and $y_p(x)$ is a particular solution of the equation, then the most general solution has the form $y(x) = y_c(x) + y_p(x)$.

An *initial value problem* consists of a linear ordinary differential equation of order n together with a set of n initial conditions of the form

$$y(x_0) = y_0, \qquad y'(x_0) = y_1, \ldots, y^{(n-1)}(x_0) = y_{n-1},$$

where $x_0 \in I$.

On the other hand, a *boundary value problem* consists of a linear ordinary differential equation of order n together with n boundary conditions that specify the values of $y(x)$ and/or its derivatives at the endpoints of the interval $I = [a,b]$. For example, if $n = 2$, then the boundary conditions may be especially simple, such as $y(a) = 0$ and $y(b) = 0$, or even more complicated, such as $\alpha y(a) + \beta y'(a) = 0$ and $\gamma y(b) + \delta y'(b) = 0$. If each boundary condition involves

only one endpoint, then the conditions are said to be *separated*. The most general boundary conditions consist of n linearly independent equations of the form

$$\sum_{n=0}^{n-1} \left(\alpha_{ik} y^{(k)}(a) + \beta_{ik} y^{(k)}(b) \right) = 0.$$

- *Integrating derivatives*: Suppose that $y(x)$ is an n-times continuously differentiable function on an interval containing the initial point $x_0 = 0$. If we set $\phi(x) = y^{(n)}(x)$, then every derivative of $y(x)$ of lower order can be expressed in terms of a single integral in which $\phi(x)$ appears in the integrand. For example, if $n = 3$, then a single integration yields

$$\begin{aligned} y''(x) &= \int_0^x y'''(t)\,\mathrm{d}t + y''(0) \\ &= \int_0^x \phi(t)\,\mathrm{d}t + y''(0). \end{aligned}$$

Another integration, followed by a change in the order of integration, yields

$$\begin{aligned} y'(x) &= \int_0^x y''(t)\,\mathrm{d}t + y'(0) \\ &= \int_0^x \left(\int_0^t \phi(s)\,\mathrm{d}s \right) \mathrm{d}t + \int_0^x y''(0)\,\mathrm{d}t + y'(0) \\ &= \int_0^x \left(\int_s^x \mathrm{d}t \right) \phi(s)\,\mathrm{d}s + y''(0)\,x + y'(0) \\ &= \int_0^x (x-s)\,\phi(s)\,\mathrm{d}s + y''(0)\,x + y'(0). \end{aligned}$$

Yet another such integration yields

$$\begin{aligned} y(x) &= \int_0^x \left(\int_0^t (t-s)\,\phi(s)\,\mathrm{d}s \right) \mathrm{d}t + \int_0^x \left(y''(0)\,t + y'(0) \right) \mathrm{d}t + y(0) \\ &= \int_0^x \left(\int_s^x (t-s)\,\mathrm{d}t \right) \phi(s)\,\mathrm{d}s + \tfrac{1}{2!}\,y''(0)\,x^2 + y'(0)\,x + y(0) \\ &= \int_0^x \frac{1}{2!}\,(x-s)^2\,\phi(s)\,\mathrm{d}s + \frac{1}{2!}\,y''(0)\,x^2 + y'(0)\,x + y(0). \end{aligned}$$

Note that the polynomial on the right-hand side of this last equation is exactly the Taylor polynomial of degree 2 for $y(x)$ constructed from the given initial data.

We turn to the general case.

Let $T_{n-1}(x)$ denote the Taylor polynomial of degree $n-1$ for $y(x)$, i.e., let

$$T_{n-1}(x) = \sum_{m=0}^{n-1} \frac{1}{m!} y^{(m)}(0) x^j.$$

If we set $\phi(x) = y^{(n)}(x)$, then after n integrations, we obtain the integral representation

$$y(x) = \int_0^x \frac{1}{(n-1)!} (x-t)^{n-1} \phi(t)\, \mathrm{d}t + T_{n-1}(x). \tag{5.1}$$

After one differentiation, we obtain

$$y'(x) = \int_0^x \frac{1}{(n-2)!} (x-t)^{n-2} \phi(t)\, \mathrm{d}t + T'_{n-1}(x).$$

In general, the k-th derivative of $y(x)$ has the representation

$$y^{(k)}(x) = \int_0^x \frac{1}{(n-k-1)!} (x-t)^{n-k-1} \phi(t)\, \mathrm{d}t + T^{(k)}_{n-1}(x). \tag{5.2}$$

Eventually, we have

$$y^{(n-2)}(x) = \int_0^x (x-t)\, \phi(t)\, \mathrm{d}t + y^{(n-1)}(0) x + y^{(n-2)}(0)$$

and

$$y^{(n-1)}(x) = \int_0^x \phi(t)\, \mathrm{d}t + y^{(n-1)}(0).$$

In a similar manner, if we are given the general differential equation

$$y^{(n)}(x) = F[x, y(x)]$$

together with n initial conditions, then we obtain the converted integral equation

$$y(x) = \int_0^x \frac{1}{(n-1)!} (x-t)^{n-1} F[t, y(t)]\, \mathrm{d}t + T_{n-1}(x). \tag{5.3}$$

The integral formulas prescribed above will be used in Sects. 5.2 and 5.3 to convert linear ordinary differential equations into both Fredholm integral equations and Volterra integral equations.

- *First-order linear systems of ordinary differential equations*:

A *first-order linear system of differential equations* has the (vector) form

$$\mathbf{v}'(x) = \mathbf{A}(x)\,\mathbf{v}(x) + \mathbf{g}(x), \tag{5.4}$$

where $\mathbf{A}(x)$ is an $n \times n$ matrix whose entries $a_{ij}(x)$ are continuous functions on some interval $I = (a, b)$. If the $n \times 1$ continuous vector $\mathbf{g}(x)$ vanishes identically on I, then the system is called *homogeneous*; if $\mathbf{g}(x)$ does not vanish identically on I, then the system is called *nonhomogeneous*. If an initial condition of the form $\mathbf{v}(x_0) = \mathbf{v}_0$ is specified, where $x_0 \in I$, then the solution to the system is unique.

The general solution $\mathbf{v}(x)$ to the linear system is always a continuously differentiable $n \times 1$ vector on I and can be decomposed into two distinct parts as $\mathbf{v}(x) = \mathbf{v}_c(x) + \mathbf{v}_p(x)$. Here, $\mathbf{v}_c(x)$ is the most general solution to the homogeneous system, and $\mathbf{v}_p(x)$ is a particular solution of the nonhomogeneous system. The vector $\mathbf{v}_c(x)$ can be expressed as the linear combination

$$\mathbf{v}_c(x) = c_1\,\mathbf{v}^{(1)}(x) + \cdots + c_n\,\mathbf{v}^{(n)}(x),$$

where each of the n vectors $\mathbf{v}^{(k)}(x)$ is a solution of the homogeneous system, and the entire set $\{\mathbf{v}^{(1)}(x), \ldots, \mathbf{v}^{(n)}(x)\}$ is linearly independent.

The invertible $n \times n$ matrix

$$\Psi(x) = (\mathbf{v}^{(1)}(x), \ldots, \mathbf{v}^{(n)}(x))$$

whose columns are the linearly independent solutions of the homogeneous system, is called the *fundamental matrix* of the system. In terms of $\Psi(x)$, we can express $\mathbf{v}_c(x)$ in the elegant form

$$\mathbf{v}_c(x) = \Psi(x)\,\mathbf{c}, \tag{5.5}$$

where $\mathbf{c}$ is an $n \times 1$ column matrix consisting of arbitrary constants. Since each column of $\Psi(x)$ is a solution to the homogeneous system, it follows that $\Psi(x)$ satisfies the matrix equation

$$\Psi'(x) = \mathbf{A}(x)\,\Psi(x). \tag{5.6}$$

The homogeneous system can be solved by converting it to a set of n linear differential equations of the nth order. For example, if $n = 2$, then the system takes the equivalent form

$$\begin{aligned} v_1'(x) &= a(x)\,v_1(x) + b(x)\,v_2(x) \\ v_2'(x) &= c(x)\,v_1(x) + d(x)\,v_2(x). \end{aligned} \tag{5.7}$$

If we differentiate these two equations, then, after some standard algebraic manipulations, we obtain the two decoupled second-order linear differential equations

$$v_1''(x) - \left(a(x) + d(x) + \frac{b'(x)}{b(x)}\right) v_1'(x) + \left(a(x)\,d(x) - b(x)\,c(x) + \frac{a(x)b'(x) - a'(x)b(x)}{b(x)}\right) v_1(x) = 0$$

and

$$v_2''(x) - \left(a(x) + d(x) + \frac{c'(x)}{c(x)}\right) v_2'(x) + \left(a(x)\,d(x) - b(x)\,c(x) + \frac{d(x)c'(x) - d'(x)c(x)}{c(x)}\right) v_2(x) = 0.$$

Each of these equations has two linearly independent solutions. For each solution $v_1(x)$ of the first equation, we can get the corresponding solution $v_2(x)$ from Eq. (5.7), and then form $\mathbf{v}^{(1)}(x)$, $\mathbf{v}^{(2)}(x)$, $\mathbf{v}_c(x)$, and $\Psi(x)$ from them. Alternately, we can solve the second equation for $v_2(x)$ and get the corresponding solution $v_1(x)$ from Eq. (5.7).

If $\mathbf{g}(x) \equiv \mathbf{0}$ on I and $\mathbf{v}(x_0)$ is specified, then $\mathbf{c} = \Psi^{-1}(x_0)\,\mathbf{v}(x_0)$, and the solution to the equation $\mathbf{v}'(x) = \mathbf{A}(x)\,\mathbf{v}(x)$ assumes the form

$$\mathbf{v}(x) = \Psi(x)\,\Psi^{-1}(x_0)\,\mathbf{v}(x_0).$$

The notation here can be simplified in a special case. If $\Psi(x_0) = \mathbf{I}$, the identity matrix, then $\Psi^{-1}(x_0) = \mathbf{I}$ as well. It is standard practice to let $\Phi(x)$ denote the fundamental matrix for which $\Phi(x_0) = \mathbf{I}$. In this case, we can simply write

$$\mathbf{v}(x) = \Phi(x)\,\mathbf{v}(x_0).$$

A particular solution $\mathbf{v}_p(x)$ to the nonhomogeneous problem can be found by the method of variation of parameters. By analogy to Eq. (5.5), we assume that the particular solution assumes the form

$$\mathbf{v}_p(x) = \Psi(x)\,\mathbf{u}_p(x), \tag{5.8}$$

where $\mathbf{u}_p(x)$ is initially unknown. After substitution into Eq. (5.4) and some simplification, we obtain

$$\Psi(x)\,\mathbf{u}_p'(x) = \mathbf{g}(x) \quad \text{or} \quad \mathbf{u}_p'(x) = \Psi^{-1}(x)\,\mathbf{g}(x).$$

Hence,

$$\mathbf{v}_p(x) = \Psi(x) \int_{x_0}^{x} \Psi^{-1}(t)\,\mathbf{g}(t)\,\mathrm{d}t,$$

and the general solution to Eq. (5.4) can be written in the form

$$\begin{aligned}\mathbf{v}(x) &= \mathbf{v}_c(x) + \mathbf{v}_p(x) \\ &= \Psi(x)\,\Psi^{-1}(x_0)\,\mathbf{v}(x) + \Psi(x) \int_{x_0}^{x} \Psi^{-1}(t)\,\mathbf{g}(t)\,\mathrm{d}t.\end{aligned}$$

5.2 Boundary Value Problems

5.2.1 Conversion to a Fredholm Integral Equation

Consider the boundary value problem that consists of the ordinary differential equation

$$\phi''(x) = F[x, \phi(x)],$$

where $F(x,y)$ is a continuous function of both of its variables, together with two prescribed values for $\phi(a)$ and $\phi(b)$. It was shown in Sect. 5.1 that this differential equation can be converted via two integrations to the corresponding integral form

$$\phi(x) = \int_{a}^{x} (x-t)\,F[t, \phi(t)]\,\mathrm{d}t + \phi(a) + \phi'(a)\,(x-a).$$

Since the constant $\phi'(a)$ is not prescribed, it should be eliminated from this formulation. If we set $x = b$, then we can solve for $\phi'(a)$ to obtain

$$\phi'(a) = \frac{\phi(b) - \phi(a)}{b-a} - \frac{1}{b-a} \int_{a}^{b} (b-t)\,F[t, \phi(t)]\,\mathrm{d}t.$$

In order to simplify the form of the converted integral equation, we assume without loss of generality that $a = 0$. After substituting $\phi'(0)$ into the integral equation, we obtain

$$\begin{aligned}\phi(x) = \phi(0) + \left(\frac{\phi(b) - \phi(0)}{b}\right) x \\ - \frac{x}{b} \int_{0}^{b} (b-t)\,F[t, \phi(t)]\,\mathrm{d}t + \int_{0}^{x} (x-t)\,F[t, \phi(t)]\,\mathrm{d}t.\end{aligned}$$

The first integral on the right can be decomposed to obtain

$$\begin{aligned}\phi(x) &= \phi(0) + \left(\frac{\phi(b)-\phi(0)}{b}\right) x \\ &\quad - \frac{x}{b}\int_x^b (b-t)\,F[t,\phi(t)]\,\mathrm{d}t - \frac{x}{b}\int_0^x (b-t)\,F[t,\phi(t)]\,\mathrm{d}t \\ &\quad + \int_0^x (x-t)\,F[t,\phi(t)]\,\mathrm{d}t \\ &= \phi(0) + \left(\frac{\phi(b)-\phi(0)}{b}\right) x \\ &\quad - \int_x^b \frac{x(b-t)}{b}\,F[t,\phi(t)]\,\mathrm{d}t - \int_0^x \frac{t(b-x)}{b}\,F[t,\phi(t)]\,\mathrm{d}t\end{aligned}$$

or

$$\phi(x) = \phi(0) + \left(\frac{\phi(b)-\phi(0)}{b}\right) x - \int_0^b G(x,t)\,F[t,\phi(t)]\,\mathrm{d}t,$$

where

$$G(x,t) = \begin{cases} \dfrac{t(b-x)}{b} & \text{if } 0 \le t \le x \le b \\ \dfrac{x(b-t)}{b} & \text{if } 0 \le x \le t \le b. \end{cases}$$

As a special case, the nonhomogeneous ordinary differential equation

$$\phi''(x) + \lambda\, q(x)\,\phi(x) = h(x),$$

where $q(x)$ and $h(x)$ are continuous on the interval $[0,b]$, can be converted to a Fredholm integral equation by letting $F[t,\phi(t)] = h(t) - \lambda\, q(t)\,\phi(t)$ in the above representation to obtain

$$\phi(x) = f(x) + \lambda \int_0^b G(x,t)\,q(t)\,\phi(t)\,\mathrm{d}t,$$

where

$$f(x) = \phi(0) + \frac{\phi(b)-\phi(0)}{b}\,x - \int_0^b G(x,t)\,h(t)\,\mathrm{d}t.$$

In particular, suppose that $q(x) \equiv 1$ and $h(x) \equiv 0$ on the interval $[0,b]$, and that $\phi(0) = \phi(b) = 0$. Then the boundary value problem consisting of the ordinary differential equation $\phi''(x) + \lambda\,\phi(x) = 0$ together with the boundary conditions $\phi(0) = \phi(b) = 0$ leads to the homogeneous Fredholm integral equation

$$\phi(x) = \lambda \int_0^b G(x,t)\,\phi(t)\,\mathrm{d}t.$$

We conclude that the solutions to the boundary value problem are actually the eigenfunctions of the triangular kernel $G(x,t)$. The kernel $G(x,t)$ is called the *Green's function* of the second derivative operator. In a more advanced text, the properties and uses of Green's functions are treated in greather depth.

5.2.2 Conversion from a Fredholm Integral Equation

Consider the Fredholm integral equation of the second kind

$$\phi(x) = f(x) + \lambda \int_0^1 K(x,t)\,\phi(t)\,\mathrm{d}t$$

with the Hermitian kernel

$$K(x,t) = \begin{cases} g(x)\,h(t) & \text{if } 0 \le x \le t \le 1 \\ g(t)\,h(x) & \text{if } 0 \le t \le x \le 1, \end{cases}$$

where $g(x)$ and $h(x)$ are continuously differentiable functions defined on the interval $[0,1]$.

Kernels of this type appear quite frequently in practice. The determination of their eigenvalues and eigenfunctions can be simplified if the integral equation is reduced to a linear differential equation of second order as follows.

After substituting $K(x,t)$ into the integral equation, we obtain

$$\phi(x) = f(x) + \lambda\,h(x)\,G(x) + \lambda\,g(x)\,H(x),$$

where we have set

$$G(x) = \int_0^x g(t)\,\phi(t)\,\mathrm{d}t \quad \text{and} \quad H(x) = \int_x^1 h(t)\,\phi(t)\,\mathrm{d}t.$$

After differentiation and simplification, we obtain

$$\phi'(x) = f'(x) + \lambda\,h'(x)\,G(x) + \lambda\,g'(x)\,H(x)$$

and

$$\phi''(x) = f''(x) + \lambda\,(h'(x)\,g(x) - g'(x)\,h(x))\,\phi(x) + \lambda\,h''(x)\,G(x) + \lambda\,g''(x)\,H(x).$$

If $G(x)$ and $H(x)$ are eliminated from these three equations, then we obtain a linear differential equation of the second order for $\phi(x)$ that can be conveniently displayed in terms of determinants as

$$\det\begin{pmatrix} \phi & g & h \\ \phi' & g' & h' \\ \phi'' & g'' & h'' \end{pmatrix} - \lambda \left[\det\begin{pmatrix} g & h \\ g' & h' \end{pmatrix}\right]^2 \phi = \det\begin{pmatrix} f & g & h \\ f' & g' & h' \\ f'' & g'' & h'' \end{pmatrix}.$$

The integral equation also implies that certain boundary conditions must necessarily be satisfied. If we set $x = 0$ in the expressions for $\phi(x)$ and $\phi'(x)$, then we obtain

$$\phi(0) = f(0) + \lambda\,g(0)\,H(0) \quad \text{and} \quad \phi'(0) = f'(0) + \lambda\,g'(0)\,H(0).$$

After eliminating $H(0)$, these boundary conditions become

$$\det\begin{pmatrix} \phi(0) & g(0) \\ \phi'(0) & g'(0) \end{pmatrix} = \det\begin{pmatrix} f(0) & g(0) \\ f'(0) & g'(0) \end{pmatrix}.$$

In a similar fashion, we obtain the additional requirement that

$$\det\begin{pmatrix} \phi(1) & h(1) \\ \phi'(1) & h'(1) \end{pmatrix} = \det\begin{pmatrix} f(1) & h(1) \\ f'(1) & h'(1) \end{pmatrix}.$$

To determine the eigenvalues and eigenfunctions of $K(x,t)$, we set $f(x) \equiv 0$, and then solve the resulting boundary value problem.

Illustrative Examples

- *Example 1*: Consider the Fredholm integral equation

$$\phi(x) = f(x) + \lambda \int_0^1 K(x,t)\,\phi(t)\,\mathrm{d}t,$$

where

$$K(x,t) = \begin{cases} x(1-t) & \text{if } 0 \le x \le t \le 1 \\ t(1-x) & \text{if } 0 \le t \le x \le 1. \end{cases}$$

We wish to determine the eigenvalues and eigenfunctions of this kernel.

By employing the method of Sect. 5.2.2, the integral equation can be converted to a linear differential equation which is easily solved.

In this example, we note that $f(x) \equiv 0$, $g(x) = x$, and $h(x) = 1 - x$. After conversion, we obtain the initial value problem consisting of the differential equation

$$\phi''(x) + \lambda\, \phi(x) = 0,$$

together with the initial conditions

$$\phi(0) = 0 \quad \text{and} \quad \phi'(0) = 0.$$

This differential equation has nontrivial solutions only if λ is positive. If we set $\lambda_n = \mu_n^2$, then the only functions that will satisfy both the linear differential equation and the accompanying boundary conditions are $\phi_n(x) = \sin(\mu_n x)$, where $\mu_n = n\pi$.

Hence, the eigenvalues of $K(x,t)$ are $\lambda_n = n^2\pi^2$, and the corresponding eigenfunctions are $\phi_n(x) = \sin(n\pi x)$. Since $\|\phi_n\|_2^2 = 1/2$, the normalized eigenfunctions of $K(x,t)$ are $\psi_n(x) = \sqrt{2}\sin(n\pi x)$.

- *Example 2*: The nontrivial solutions of the integral equation

$$\phi(x) = \lambda \int_0^1 J(x,t)\, \phi(t)\, \mathrm{d}t$$

are the eigenfunctions of the symmetric kernel

$$J(x,t) = \begin{cases} -\sqrt{xt}\, \ln t & \text{if } 0 \le x \le t \le 1 \\ -\sqrt{xt}\, \ln x & \text{if } 0 \le t \le x \le 1. \end{cases}$$

We determine the eigenvalues and the eigenfunctions of $J(x,t)$. In this example, $f(x) \equiv 0$, $g(x) = -\sqrt{x}$, and $h(x) = \sqrt{x}\ln x$. Note that we have defined $h(0) = 0$ by continuous extension, after having applied l'Hôpital's rule. Since $J(x,t)$ is continuous on the square $Q(0,1)$, its eigenfunctions are continuous on the interval $[0,1]$.

After conversion, we obtain the linear differential equation

$$x^2\, \phi''(x) + \left(\lambda\, x^2 + \tfrac{1}{4}\right)\, \phi(x) = 0$$

together with the initial condition $\phi(1) = 0$. Note that his equation has a regular singular point at $x = 0$. Every solution of this equation is a linear combination of

two linearly independent solutions. In order to determine its solutions, we appeal to a known result within the theory of Bessel functions. The solutions of the differential equation

$$x^2 y''(x) + \left(\alpha^2\beta^2 x^{2\beta} + \tfrac{1}{4} - \nu^2\beta^2\right) y(x) = 0$$

on the interval $(0,+\infty)$ assume the form $y = x^{1/2} f(\alpha x^{\beta})$, where $f(\zeta)$ is a solution of the Bessel equation of order ν.

In the present situation, $\nu = 0$, $\beta = 1$, and $\lambda = \alpha^2$. Thus, the solution to the differential equation of interest assumes the form

$$\phi(x) = x^{1/2}\left(k_1 J_0(\alpha x) + k_2 Y_0(\alpha x)\right),$$

where $J_0(x)$ and $Y_0(x)$ are the Bessel functions of order 0. However, since $Y_0(x)$ has a logarithmic singularity at $x = 0$, it is not a part of any eigenfunction that is continuous on the interval $[0,1]$. Also, the initial condition implies that $\phi(1) = k_1 J_0(\alpha) = 0$. Hence, each α is a zero of $J_0(x)$.

It follows that the eigenvalues of the kernel are the squares of the zeroes of $J_0(x)$. If we denote the zeroes of $J_0(x)$ by $\{\alpha_n\}$, then the eigenvalues of the kernel $J(x,t)$ are $\lambda_n = \alpha_n^2$, and the corresponding eigenfunctions are

$$\phi_n(x) = x^{1/2} J_0(\alpha_n x).$$

These eigenfunctions are orthogonal since it is well known that

$$\int_0^1 \phi_n(x)\,\phi_m(x)\,\mathrm{d}x = \int_0^1 x J_0(\alpha_n x)\,J_0(\alpha_m x)\,\mathrm{d}x = 0$$

whenever $n \neq m$. Furthermore, since

$$\|\phi_n\|_2^2 = \int_0^1 x J_0^2(\alpha_n x)\,\mathrm{d}x = \frac{1}{2}\left(J_0'(\alpha_n)\right)^2 = \frac{1}{2} J_1^2(\alpha_n),$$

the functions

$$\psi_n(x) = \frac{\sqrt{2}\,x^{1/2} J_0(\alpha_n x)}{J_1(\alpha_n)}$$

are the orthonormalized eigenfunctions of the kernel $J(x,t)$.

Since the eigenvalues λ_n of $J(x,t)$ are positive, it is a positive kernel. Hence, Theorem 3.4.2 (Mercer's theorem) allows us to conclude that $J(x,t)$ has the uniformly convergent representation

$$J(x,t) = \sum_{n=1}^{\infty} \frac{2\sqrt{xt}\,J_0(\alpha_n x)\,J_0(\alpha_n t)}{\alpha_n^2 J_1^2(\alpha_n)}$$

on the square $Q(0,1)$. In Exercise 4 of Sect. 3.5, it was shown that the first trace A_1 of $J(x,t)$ is equal to $1/4$ and that the second trace A_2 is equal to $1/32$. Now that the eigenvalues of $J(x,t)$ have been determined, we can conclude that

$$\sum_{n=1}^{\infty} \frac{1}{\alpha_n^2} = \frac{1}{4}$$

and

$$\sum_{n=1}^{\infty} \frac{1}{\alpha_n^4} = \frac{1}{32}.$$

Section 5.2 Exercises

1. Consider the following boundary value problem consisting of the ordinary differential equation

$$\phi''(x) + \pi^2 \phi(x) = \cos(\pi x)$$

together with the boundary conditions $\phi(0) = 1$ and $\phi(\frac{1}{2}) = \frac{1}{2\pi}$.

(a) Show that the unique solution to this boundary value problem is

$$\phi(x) = \cos(\pi x) + \frac{1}{4\pi} \sin(\pi x) + \frac{1}{2\pi} x \sin(\pi x).$$

(b) By using the method illustrated in Sect. 5.2.1, convert this differential equation into a Fredholm integral equation of the second kind.

(c) Show that the solution to the differential equation is also the solution to your converted integral equation.

Hint: You will need to evaluate the integral

$$\int_0^{1/2} K(x,t) \cos(\pi t)\, dt = (1-2x) \int_0^x t \cos(\pi t)\, dt + x \int_x^{1/2} (1-2t) \cos(\pi t)\, dt.$$

2. In this exercise, we generalize the results of Example 2. The nontrivial solutions of the integral equation

$$\phi(x) = \lambda \int_0^1 J(x,t;\alpha)\, \phi(t)\, dt$$

are the eigenfunctions of the kernel

$$J(x,t;\alpha) = \begin{cases} -(xt)^{\alpha} \ln t & \text{if } 0 \le x \le t \le 1 \\ -(xt)^{\alpha} \ln x & \text{if } 0 \le t \le x \le 1, \end{cases}$$

where $\alpha > 0$. We determine the eigenvalues and the eigenfunctions of $J(x,t;\alpha)$.

(a) Let $g(x) = -x^{\alpha}$ and $h(x) = x^{\alpha} \ln x$. Apply the conversion technique to show that $\phi(x)$ satisfies the differential equation

$$x^2 \phi''(x) + (1-2\alpha) x\phi'(x) + \left(\alpha^2 + \lambda x^{2\alpha+1}\right) \phi(x) = 0.$$

(b) Let $\phi(x) = x^{\alpha} \sigma(x)$ and show that $\sigma(x)$ satisfies the differential equation

$$x\sigma''(x) + \sigma'(x) + \lambda x^{2\alpha} \sigma(x) = 0.$$

(c) Show that

$$\sigma(x) = J_0\left(\tfrac{2\sqrt{\lambda}}{2\alpha+1} x^{(2\alpha+1)/2}\right)$$

is a solution of the differential equation in part (b), where J_0 is the Bessel function of the first kind of order zero. Thus, the solution assumes the form

$$\phi(x) = x^{\alpha} J_0\left(\tfrac{2\sqrt{\lambda}}{2\alpha+1} x^{(2\alpha+1)/2}\right).$$

(d) With $\phi(1) = 0$, deduce that the eigenvalues of $J(x,t;\alpha)$ are

$$\lambda_n = \alpha_n^2 \left(\tfrac{2\alpha+1}{2}\right)^2,$$

where $J_0(\alpha_n) = 0$, and therefore that

$$\phi_n(x) = x^{\alpha} J_0\left(\alpha_n x^{(2\alpha+1)/2}\right).$$

Note that if $\alpha = \frac{1}{2}$, then these results agree with the results in Example 2.

(e) Show that the set $\{\phi_n(x)\}_{n=1}^{\infty}$ is mutually orthogonal, i.e., that

$$\int_0^1 \phi_n(x)\, \phi_m(x)\, \mathrm{d}x = 0$$

if $n \neq m$.

Hint: Introduce the change of variables

$$u = x^{(2\alpha+1)/2},$$

for which it can be shown that

$$x^{2\alpha}\,\mathrm{d}x = \tfrac{2}{2\alpha+1}\,u\,\mathrm{d}u.$$

(f) Show that

$$\|\phi_n\|_2^2 = \int_0^1 x^{2\alpha} J_0^2\left(\alpha_n x^{(2\alpha+1)/2}\right)\mathrm{d}x = \tfrac{1}{2\alpha+1} J_1^2(\alpha_n).$$

(g) Use parts (e) and (f) to produce the infinite bilinear expansion for $J(x,t;\alpha)$.
(h) Show that the trace $A_1(\alpha)$ of $J(x,t;\alpha)$ is given by

$$A_1(\alpha) = \frac{1}{(2\alpha+1)^2}.$$

5.3 Initial Value Problems

5.3.1 Conversion to a Volterra Integral Equation

Consider the initial value problem consisting of the linear ordinary differential equation of order n

$$y^{(n)}(x) + a_1(x)y^{(n-1)}(x) + \cdots + a_{n-1}y'(x) + a_n(x)y(x) = g(x),$$

where $g(x)$ and the coefficient functions $a_i(x)$ are assumed to be continuous on some interval $I = (a,b)$, together with the initial conditions

$$y(x_0) = y_0, \qquad y'(x_0) = y_1, \ldots, y^{(n-1)}(x_0) = y_{n-1},$$

where $x_0 \in I$. It is possible to convert an initial value problem of this form into a linear Volterra integral equation of the second kind. Without loss of generality, we assume that $x_0 = 0$.

If we now substitute the integral representation (5.2) for $y^{(k)}(x)$ that was determined in Sect. 5.1 into the differential equation above, then we obtain

$$\phi(x) + \sum_{k=0}^{n-1} a_{n-k}(x)\left(\int_0^x \frac{1}{(n-k-1)!}(x-t)^{n-k-1}\,\phi(t)\,\mathrm{d}t + T_{n-1}^{(k)}(x)\right) = g(x).$$

Upon rearrangement, we obtain

$$\phi(x) + \int_0^x \left(\sum_{k=0}^{n-1} \frac{1}{(n-k-1)!}\,a_{n-k}(x)\,(x-t)^{n-k-1}\right)\phi(t)\,\mathrm{d}t$$

$$= g(x) - \sum_{k=0}^{n-1} a_{n-k}(x)\,T_{n-1}^{(k)}(x).$$

The converted equation is clearly a Volterra integral equation of the second kind of the form

$$\phi(x) = f(x) + \int_0^x K(x,t)\,\phi(t)\,\mathrm{d}t,$$

where

$$K(x,t) = -\sum_{k=0}^{n-1} \frac{1}{(n-k-1)!}\,a_{n-k}(x)\,(x-t)^{n-k-1}$$

and

$$f(x) = g(x) - \sum_{k=0}^{n-1} a_{n-k}(x)\,T_{n-1}^{(k)}(x).$$

Note that if the coefficient functions $a_{n-k}(x)$ are constants, then the kernel $K(x,t)$ is a convolution kernel. Equations of this type were solved in Sect. 4.3 by employing the Laplace transform.

After the solution $\phi(x)$ to the Volterra integral equation has been found, it can be substituted directly into the representation (5.1) to obtain an integral representation for the solution to the given differential equation.

5.3.2 *Conversion from a Volterra Integral Equation*

Consider the Volterra integral equation of the second kind

$$\phi(x) = f(x) + \lambda \int_0^x K(x,t)\,\phi(t)\,\mathrm{d}t$$

with the separable kernel

$$K(x,t) = \sum_{i=1}^{n} a_i(x)\,b_i(t).$$

Upon substitution, we obtain

$$\phi(x) = f(x) + \lambda \sum_{i=1}^{n} a_i(x)\,v_i(x),$$

where

$$v_i(x) = \int_0^x b_i(t)\,\phi(t)\,\mathrm{d}t.$$

If we differentiate these integrals and then substitute the expression for $\phi(x)$, we obtain the first-order linear system of differential equations

$$v_i'(x) = b_i(x)\, f(x) + \lambda \sum_{j=1}^{n} a_j(x) b_i(x)\, v_j(x),$$

where $i = 1, \ldots, n$. In vector form, the linear system becomes

$$\mathbf{v}'(x) = \lambda\, \mathbf{M}(x)\, \mathbf{v}(x) + f(x)\, \mathbf{b}(x), \tag{5.9}$$

where

$$\mathbf{v}(x) = \begin{pmatrix} v_1(x) \\ \vdots \\ v_n(x) \end{pmatrix}, \qquad \mathbf{a}(x) = \begin{pmatrix} a_1(x) \\ \vdots \\ a_n(x) \end{pmatrix}, \qquad \mathbf{b}(x) = \begin{pmatrix} b_1(x) \\ \vdots \\ b_n(x) \end{pmatrix},$$

and $\mathbf{M}(x) = \mathbf{b}(x) \cdot \mathbf{a}^{\mathrm{T}}(x)$. Since $v_i(0) = 0$ for each $i = 1, \ldots, n$, the initial vector condition is $\mathbf{v}(0) = \mathbf{0}$. For example, if $n = 2$, then this linear system takes the expanded form

$$\frac{\mathrm{d}}{\mathrm{d}x} \begin{pmatrix} v_1(x) \\ v_2(x) \end{pmatrix} = \lambda \begin{pmatrix} a_1(x)\, b_1(x) & a_2(x)\, b_1(x) \\ a_1(x)\, b_2(x) & a_2(x)\, b_2(x) \end{pmatrix} \begin{pmatrix} v_1(x) \\ v_2(x) \end{pmatrix} + f(x) \begin{pmatrix} b_1(x) \\ b_2(x) \end{pmatrix}.$$

After having solved this linear system for $\mathbf{v}(x)$ as indicated in Sect. 5.1, we can write the solution to the integral equation in the vector form

$$\phi(x) = f(x) + \lambda\, \mathbf{a}(x) \cdot \mathbf{v}(x).$$

Illustrative Examples

- *Example 1*: Convert the initial value problem

  $$y''(x) + x y'(x) + y(x) = 0$$

  with the given initial conditions $y(0) = 1$ and $y'(0) = 0$ to a Volterra integral equation of the second kind.

  It is easy to solve this differential equation, since it can be recast in the form $y''(x) + (x y(x))' = 0$. After two integrations, we find that its general solution is

  $$y(x) = c_1\, \mathrm{e}^{-x^2/2} + c_2\, \mathrm{e}^{-x^2/2} \int_0^x \mathrm{e}^{t^2/2} \mathrm{d}t.$$

  The given initial conditions imply that $c_1 = 1$ and that $c_2 = 0$. Hence, $y(x) = \mathrm{e}^{-x^2/2}$.

If we set

$$y''(x) = \phi(x),$$

$$y'(x) = \int_0^x \phi(t)\,\mathrm{d}t + y'(0),$$

and

$$y(x) = \int_0^x (x-t)\,\phi(t)\,\mathrm{d}t + y'(0)\,x + y(0),$$

according to the discussion in Sect. 5.1, then we obtain the integral equation

$$\phi(x) = -1 - \int_0^x (2x-t)\,\phi(t)\,\mathrm{d}t.$$

The solution to this equation is

$$\phi(x) = y''(x) = (\mathrm{e}^{-x^2/2})'' = (x^2-1)\,\mathrm{e}^{-x^2/2},$$

as can be easily verified.

- *Example 2*: The Volterra integral equation of the second kind

$$\phi(x) = f(x) + \lambda \int_0^x a(x)\,b(t)\,\phi(t)\,\mathrm{d}t$$

is easily converted to a linear differential equation of the first order. If we set

$$y(x) = \int_0^x b(t)\,\phi(t)\,\mathrm{d}t,$$

then

$$\begin{aligned} y'(x) &= b(x)\,\phi(x) \\ &= b(x)\,[f(x) + \lambda\,a(x)\,y(x)], \end{aligned}$$

or

$$y'(x) - \lambda\,a(x)\,b(x)\,y(x) = b(x)\,f(x),$$

together with the initial condition $y(0) = 0$.

- *Example 3*: Consider the linear Volterra integral equation

$$\phi(x) = \frac{x^3}{6} + \lambda \int_0^x K(x,t)\,\phi(t)\,\mathrm{d}t. \tag{5.10}$$

where $K(x,t) = x - t$ and $\lambda = -1$. As shown in Sect. 4.3, Volterra equations with difference kernels can be solved by employing the Laplace transform. After a short calculation, we find that the solution to this equation is $\phi(x) = x - \sin x$.

Here, we illustrate a second method of finding the solution to this integral equation that is based upon conversion to a linear system of equations, as discussed in Sect. 5.3.2. With this method, the solution assumes the form

$$\phi(x) = f(x) - \mathbf{a}(x) \cdot \mathbf{v}(x),$$

where $\mathbf{v}(x)$ is the solution of the converted first-order linear system

$$\mathbf{v}'(x) = \lambda\, \mathbf{M}(x)\, \mathbf{v}(x) + f(x)\, \mathbf{b}(x) \tag{5.11}$$

with $\mathbf{v}(0) = \mathbf{0}$.

Since $a_1(x) = x$, $b_1(t) = 1$, $a_2(x) = -1$, and $b_2(t) = t$, the linear system assumes the equivalent form,

$$\frac{\mathrm{d}}{\mathrm{d}x} \begin{pmatrix} v_1(x) \\ v_2(x) \end{pmatrix} = - \begin{pmatrix} x & -1 \\ x^2 & -x \end{pmatrix} \begin{pmatrix} v_1(x) \\ v_2(x) \end{pmatrix} + \frac{x^3}{6} \begin{pmatrix} 1 \\ x \end{pmatrix}.$$

We solve this system by following the method outlined in Sect. 5.1. Our first task is to determine $\mathbf{v}_c(x)$. After decoupling this system, we find that $v_1(x)$ satisfies the equation

$$v_1''(x) + v_1(x) = 0,$$

whose general solution is

$$v_1(x) = c_1 \cos x + c_2 \sin x,$$

and that $v_2(x)$ satisfies the equation

$$x\, v_2''(x) - 2\, v_2'(x) + x\, v_2(x) = 0,$$

whose general solution is

$$v_2(x) = c_1\, (\cos x + x \sin x) + c_2\, (x \cos x - \sin x).$$

Since $v_2(x) = v_1'(x) + x\, v_1(x)$, it follows that

$$\mathbf{v}^{(1)}(x) = \begin{pmatrix} \cos x \\ x \cos x - \sin x \end{pmatrix} \quad \text{and} \quad \mathbf{v}^{(2)}(x) = \begin{pmatrix} \sin x \\ x \sin x + \cos x \end{pmatrix}.$$

The general solution $\mathbf{v}_c(x) = c_1\, \mathbf{v}^{(1)}(x) + c_2\, \mathbf{v}^{(2)}(x)$ is prescribed in terms of these vector functions.

The fundamental matrix for this linear system is

$$\Phi(x) = \begin{pmatrix} \cos x & \sin x \\ x\cos x - \sin x & x\sin x + \cos x \end{pmatrix}.$$

Note that $\Phi(0) = I$. With $\det\Phi(x) = 1$, we have

$$\Phi^{-1}(x) = \begin{pmatrix} x\sin x + \cos x & -\sin x \\ \sin x - x\cos x & \cos x \end{pmatrix}.$$

To determine a particular solution $\mathbf{v}_p(x) = \Phi(x)\,\mathbf{u}_p(x)$, we follow the procedure outlined in Sect. 5.1 to obtain

$$\mathbf{u}_p'(x) = \Phi^{-1}(x)\,\mathbf{b}(x) = \frac{1}{6}\begin{pmatrix} x^3\cos x \\ x^3\sin x \end{pmatrix}.$$

After integration, we have

$$\mathbf{u}_p(x) = \frac{1}{6}\begin{pmatrix} (3x^2-6)\cos x + (x^3-6x)\sin x \\ (6x-x^3)\cos x + (3x^2-6)\sin x \end{pmatrix}.$$

After multiplication, we have

$$\mathbf{v}_p(x) = \begin{pmatrix} \frac{1}{2}x^2 - 1 \\ \frac{1}{3}x^3 \end{pmatrix}.$$

The general solution to the linear system has the form

$$\mathbf{v}(x) = c_1\,\mathbf{v}^{(1)}(x) + c_2\,\mathbf{v}^{(2)}(x) + \mathbf{v}_p(x).$$

The initial condition $\mathbf{v}(0) = \mathbf{0}$ implies that $c_1 = 1$ and $c_2 = 0$. Thus,

$$\mathbf{v}(x) = \begin{pmatrix} \cos x + \frac{1}{2}x^2 - 1 \\ x\cos x - \sin x + \frac{1}{3}x^3 \end{pmatrix}.$$

Finally, the solution $\phi(x)$ to the Volterra integral equation becomes

$$\phi(x) = f(x) - \mathbf{a}(x)\cdot\mathbf{v}(x) = x - \sin x.$$

Section 5.3 Exercises

1. Convert each of the following Volterra integral equations into first-order linear ordinary differential equations and then determine their solutions:

   (a) $\phi(x) = a + bx + cx^2 + \lambda \int_0^x \phi(t)\,\mathrm{d}t$

   (b) $\phi(x) = a\mathrm{e}^{bx} + c + \lambda \int_0^x \phi(t)\,\mathrm{d}t$

   (c) $\phi(x) = a + bx + \lambda \int_0^x xt\,\phi(t)\,\mathrm{d}t$

   (d) $\phi(x) = a\mathrm{e}^{x} + b + \lambda \int_0^x \mathrm{e}^{x-t}\,\phi(t)\,\mathrm{d}t$

   (e) $\phi(x) = a\sec x + b\tan x + \lambda \int_0^x \sec x \cos t\,\phi(t)\,\mathrm{d}t$

   Answer: $\phi(x) = (a + \frac{b}{1+\lambda^2})\,\mathrm{e}^{\lambda x}\sec x + \frac{b}{1+\lambda^2}(\tan x - \lambda)$.

   (f) $\phi(x) = ax^2 + bx + \lambda \int_0^x \dfrac{ax+b}{at+b}\,\phi(t)\,\mathrm{d}t$

   Answer: $\phi(x) = \dfrac{1}{\lambda}(ax+b)(\mathrm{e}^{\lambda x} - 1)$.

2. (a) Show that the unique solution to the initial value problem consisting of the differential equation

$$y''(x) + xy'(x) + xy(x) = 0$$

   together with the initial conditions $y(0) = y'(0) = 1$ is given by $y(x) = \mathrm{e}^{x - x^2/2}$.

   (b) Show that the differential equation in part (a) can be converted into the Volterra integral equation

$$\phi(x) = -x^2 - 2x - \int_0^x \left(x^2 + x - xt\right)\phi(t)\,\mathrm{d}t$$

   and verify that its solution is $\phi(x) = y''(x) = x(x-2)\,\mathrm{e}^{x - x^2/2}$.

3. Two linearly independent series solutions of Airy's differential equation

$$w''(x) - xw(x) = 0$$

   are given by

$$w_1(x) = \sum_{n=0}^{\infty} 3^n \left(\tfrac{1}{3}\right)_n \frac{x^{3n}}{(3n)!} = 1 + \frac{1}{3!}x^3 + \cdots$$

   and

$$w_2(x) = \sum_{n=0}^{\infty} 3^n \left(\tfrac{2}{3}\right)_n \frac{x^{3n+1}}{(3n+1)!} = x + \frac{2}{4!}x^4 + \cdots.$$

In terms of these series, the solutions to Airy's equation are expressed in the forms

$$Ai(x) = c_1 w_1(x) - c_2 w_2(x) \quad \text{and} \quad Bi(x) = \sqrt{3}\,(c_1 w_1(x) + c_2 w_2(x))\,,$$

where

$$c_1 = Ai(0) = \frac{Bi(0)}{\sqrt{3}} = \frac{1}{3^{2/3}\,\Gamma(\frac{2}{3})}$$

and

$$c_2 = -Ai'(0) = \frac{Bi'(0)}{\sqrt{3}} = \frac{1}{3^{1/3}\,\Gamma(\frac{1}{3})}.$$

(a) Convert Airy's differential equation into Volterra integral equations, one satisfied by $w_1(x)$ and another by $w_2(x)$.

(b) Show that two linearly independent solutions of the linear differential equation

$$y''(x) + y'(x) + x\,y(x) = 0$$

are

$$y_1(x) = \mathrm{e}^{-x/2} Ai\left(\tfrac{1}{4} - x\right) \quad \text{and} \quad y_2(x) = \mathrm{e}^{-x/2} Bi\left(\tfrac{1}{4} - x\right).$$

(c) Convert the differential equation in part (b) into a Volterra integral equation of the second kind.

4. The classical Hermite differential equation is given by

$$y''(x) - 2x\,y'(x) + 2n\,y(x) = 0.$$

If n is a nonnegative integer, then one solution $H_n(x)$ of this equation is a polynomial of degree n. These polynomials appear as the coefficients in the generating function

$$\exp\{2xt - t^2\} = \sum_{k=0}^{\infty} \frac{1}{k!} H_k(x)\,t^k.$$

(a) Set $x = 0$ in the generating function to determine the value of $H_k(0)$ for all $k \geq 0$.

(b) Differentiate the generating function with respect to x to show that

$$H_k'(x) = 2k\,H_{k-1}(x)$$

for all $k \geq 1$. Use this result to determine the value of $H_k'(0)$ for all $k \geq 1$.

(c) For $n \geq 1$, convert Hermite's differential equation together with the values of $H_n(0)$ and $H_n'(0)$ to a Volterra integral equation of the second kind.

5.4 Integrodifferential Equations

An integrodifferential equation can be viewed as a hybrid of an integral equation and an ordinary or partial differential equation, in the sense that it involves both the integral and the derivative(s) of the unknown function.

Since the variety of integrodifferential equations is essentially infinite, no satisfactory classification of them is in existence. Consequently, the methods available for determining their solutions vary considerably and often involve some ad hoc creativity. In some situations, a solution can be obtained directly; in others, a solution can be obtained after converting the integrodifferential equation into either a differential equation or an integral equation.

Qualitative and quantitative theoretical investigations of integrodifferential equations are generally concerned with the existence and uniqueness of their solutions, along with an examination of their analytic properties. These include periodicity, boundedness, stability analysis, and continuous dependence upon parameters. However, since these studies would take us too far afield from the introductory nature of this textbook, we only consider here a few types of solvable equations.

Illustrative Examples

- *Example 1*: Consider the integrodifferential equation

$$\phi''(x) + \lambda\,\phi(x) = \int_0^1 \phi(t)\,\mathrm{d}t.$$

This equation can be considered to be of Fredholm type, since the upper and lower limits of integration are constants. Since the right-hand side of this equation is constant, any solution to it must be found among the solutions to the third-order linear differential equation

$$\phi'''(x) + \lambda\,\phi'(x) = 0,$$

which is immediately obtained by differentiation. Regardless of the value of λ, this differential equation will always have three linearly independent solutions.

We consider three cases:

Case I: $\lambda = 0$.

In this case, the most general solution to the differential equation assumes the form

$$\phi(x) = \phi(0) + \phi'(0)\,x + \tfrac{1}{2}\,\phi''(0)\,x^2.$$

However, if we substitute this solution into the integrodifferential equation, then the necessary condition

$$\phi''(0) = \int_0^1 \phi(t)\,\mathrm{d}t = \phi(0) + \tfrac{1}{2}\,\phi'(0) + \tfrac{1}{6}\,\phi''(0),$$

or equivalently,

$$\phi''(0) = \tfrac{6}{5}\,\phi(0) + \tfrac{3}{5}\,\phi'(0)$$

must be satisfied for a solution to exist. Two of the initial conditions can be chosen arbitrarily. For example, if we substitute the above value for $\phi''(0)$, we obtain

$$\phi(x) = \phi(0)\,\left(1 + \tfrac{3}{5}\,x^2\right) + \phi'(0)\,\left(x + \tfrac{3}{10}\,x^2\right).$$

Case II: $\lambda = +\sigma^2$ with $\sigma > 0$.

In this case, the most general solution to the differential equation assumes the form

$$\phi(x) = \phi(0) + \frac{1}{\sigma}\,\phi'(0)\,\sin(\sigma x) + \frac{1}{\sigma^2}\,\phi''(0)\,(1 - \cos(\sigma x)).$$

However, if we substitute this solution into the integrodifferential equation, then the necessary condition

$$\phi''(0)\,\left(1 - \frac{1}{\sigma^2} + \frac{\sin\sigma}{\sigma^3}\right) = \phi(0)\,(1 - \sigma^2) + \frac{1}{\sigma^2}\,\phi'(0)\,(1 - \cos\sigma)$$

must be satisfied for a solution to exist. Special care should be exercised if $\sigma = 1$ or $\sigma = n\pi$.

Case III: $\lambda = -\sigma^2$ with $\sigma > 0$.

In this case, the most general solution to the differential equation assumes the form

$$\phi(x) = \phi(0) + +\phi'(0)\,\left(\frac{\sinh(\sigma x)}{\sigma}\right) + 2\phi''(0)\,\left(\frac{\sinh(\sigma x/2)}{\sigma}\right)^2.$$

However, if we substitute this solution into the integrodifferential equation, then the necessary condition

$$\phi''(0)\,\left(1 + \frac{1}{\sigma^2} - \frac{\sinh\sigma}{\sigma^3}\right) = \phi(0)\,(1 + \sigma^2) + \frac{1}{\sigma^2}\,\phi'(0)\,(\cosh\sigma - 1)$$

must be satisfied for a solution to exist.

- *Example 2*: Consider the Fredholm-type integrodifferential equation

$$\phi(x) = x^3 + \lambda \int_0^1 x^3 t^3 \, \phi'(t) \, \mathrm{d}t.$$

This integral equation can be solved directly. If we set

$$A = \int_0^1 t^3 \, \phi'(t) \, \mathrm{d}t,$$

then the solution assumes the form $\phi(x) = (1 + A\lambda) x^3$. After substituting this representation into the integral equation and then simplifying the result, we find that $A = 1/(2-\lambda)$. Consequently, if $\lambda \neq 2$, then

$$\phi(x) = \frac{2x^3}{2-\lambda}.$$

- *Example 3*: Consider the integrodifferential equation

$$\phi''(x) - 3\,\phi(x) = \int_0^x 2\,\phi(t) \, \mathrm{d}t.$$

This equation can be considered to be of Volterra type, since the upper limit of integration is variable. Any solution to it must be found among the solutions to the third-order linear differential equation

$$\phi'''(x) - 3\,\phi'(x) - 2\,\phi(x) = 0,$$

which is directly obtained by differentiation. The most general solution to this differential equation can be written in the form

$$\begin{aligned}\phi(x) = {} & \frac{1}{15}\left(\phi(0) + 3\,\phi'(0) + 2\,\phi''(0)\right) \mathrm{e}^{2x} \\ & + \frac{1}{15}\left(8\,\phi(0) - 6\,\phi'(0) + \phi''(0)\right) \mathrm{e}^{-x} \\ & + \frac{1}{15}\left(6\,\phi(0) + 3\,\phi'(0) - 3\,\phi''(0)\right) x\mathrm{e}^{-x}.\end{aligned}$$

However, if we substitute this solution into the integrodifferential equation, then we find that the necessary condition $2\,\phi''(0) = 9\,\phi(0) - 3\,\phi'(0)$ must be satisfied for a solution to exist. If we substitute this value for $\phi''(0)$, we obtain

$$\phi(x) = \frac{1}{6}\,\phi(0)\,\left(4\mathrm{e}^{2x} + 5\mathrm{e}^{-x} - 3x\mathrm{e}^{-x}\right) - \frac{1}{2}\,\phi'(0)\,\left(\mathrm{e}^{-x} - x\mathrm{e}^{-x}\right),$$

which is the most general solution to the integrodifferential equation.

- *Example 4*: Consider the integrodifferential equation

$$\phi'(x) = \cos x + 2\int_0^x \sin(x-t)\,\phi(t)\,\mathrm{d}t$$

together with the initial condition $\phi(0) = 1$. This Volterra-type equation can be solved by employing the method of Laplace transforms that was introduced in Sect. 4.3, since the kernel is of convolution type. After taking the Laplace transform of both sides of this equation and then simplifying the result, we obtain

$$\mathcal{L}\{\phi\} = \frac{s^2+s+1}{s^3+s^2-2} = \frac{3}{5}\,\frac{1}{s-1} + \frac{1+2s}{s^2+2s+2}.$$

Hence,

$$\phi(x) = \frac{3}{5}\,\mathrm{e}^{x} + \frac{2}{5}\,\mathrm{e}^{-x}\cos x - \frac{1}{5}\,\mathrm{e}^{-x}\sin x.$$

Section 5.4 Exercises

1. Consider the Fredholm-type integrodifferential equation

$$\phi(x) = x^3 + \lambda\int_0^1 x^3 t^3\,\phi^{(n)}(t)\,\mathrm{d}t.$$

This equation was solved in *Example 2* if $n = 1$. Use the technique illustrated there to solve this equation if $n = 2$ and if $n = 3$. What happens if $n \geq 4$?

2. Consider the Fredholm-type integrodifferential equation

$$\phi^{(n)}(x) = x^3 + \lambda\int_0^1 x^3 t^3\,\phi(t)\,\mathrm{d}t.$$

Solve this equation if $n = 1$ and if $n = 2$. What happens if $n \geq 3$?

3. Let n be a positive integer. Solve the Volterra-type integrodifferential equation

$$\phi(x) = \frac{1}{n!}\,x^n + \frac{1}{n!}\int_0^x (x-t)^n\,\phi'(t)\,\mathrm{d}t$$

for $n = 1, 2$, and 3. Note that $\phi(0) = 0$. Since the kernel is of convolution type, the method introduced in Sect. 4.3 can be employed here.

4. Solve the Volterra-type integrodifferential equation

$$\int_1^x \frac{1}{\sqrt{xt}}\,\phi'(t)\,\mathrm{d}t = f(x),$$

where we assume that $f(1) = 0$ and $f(x)$ is continuously differentiable on an interval of the form $[1, b]$.

Chapter 6
Nonlinear Integral Equations

A *nonlinear integral equation* is an integral equation in which the unknown function appears in the equation in a nonlinear manner. The nonlinearity may occur either inside or outside of the integrand or simultaneously in both of these locations. It leads to an astonishing variety of new phenomena related to the characteristics of the solutions and to the methods of solution.

Since the variety of nonlinear integral equations is almost unlimited in scope, a definitive classification of them is elusive. However, one of the most common categories is that of equations assuming the form

$$H(x,\phi(x)) = \int_a^b G(x,t,\phi(t))\,\mathrm{d}t.$$

We consider equations of this general form, but we are mainly concerned with equations of two special types:

- The Fredholm type

$$\phi(x) = f(x) + \lambda \int_a^b G(x,t,\phi(t))\,\mathrm{d}t, \tag{6.1}$$

  for which we always assume that the integrand $G(x,t,\phi(t))$ is continuous either on the set $Q(a,b)\times\mathbb{R}$ or on the smaller set $Q(a,b)\times[-M,+M]$ for a suitably large value of M.
- The Volterra type

$$\phi(x) = f(x) + \lambda \int_a^x G(x,t,\phi(t))\,\mathrm{d}t, \tag{6.2}$$

  for which we always assume that the integrand $G(x,t,\phi(t))$ is continuous either on the set $T(a,b)\times\mathbb{R}$ or on the smaller set $T(a,b)\times[-M,+M]$ for a suitable value of M.

S.M. Zemyan, *The Classical Theory of Integral Equations: A Concise Treatment*,
DOI 10.1007/978-0-8176-8349-8_6, © Springer Science+Business Media, LLC 2012

In these equations, we always assume that the inhomogeneous term $f(x)$ is continuous on the interval $[a,b]$.

It is natural to ask whether solutions to nonlinear integral equations exist and, if so, whether these solutions are unique. It is also reasonable to explore the role that the parameter λ plays in the determination of these solutions. Although straightforward routines exist to determine these solutions when the integral equations assume familiar forms, we find that creative ad hoc methods are also required when they do not. Our analysis of these questions reveals that new phenomena occur here that did not occur in our previous work with linear integral equations.

When a solution cannot be determined exactly, numerical techniques can be employed to determine approximate solutions. The nonlinearity of the integral equations induces additional computational difficulties into our work.

In Sect. 6.1, we present several tools of the trade that are indispensible for the comprehension of the material in this chapter.

In Sect. 6.2, we apply a fixed point theorem originally enunciated by Banach to establish the existence and uniqueness of a solution to an integral equation of Fredholm type (6.1) for restricted values of λ, when the integrand satisfies a particular analytic condition. An analogous theorem for nonlinear integral equations of Volterra type (6.2) is also proven. Other fixed point theorems, e.g., those of Brouwer, Schauder, and Kakutani, can also be invoked to prove existence and uniqueness theorems in different settings, but we shall not pursue those here. We also briefly consider equations of Hammerstein type. These equations are interesting since their solution involves an application of the Hilbert–Schmidt theory that was introduced in Chap. 3.

In Sect. 6.3, we consider a large number of examples in order to illustrate many of the new phenomena that can and do occur in an investigation of nonlinear integral equations. There is no need to consider complicated integral equations to illustrate these phenomena. Indeed, extremely simple equations can be used to demonstrate some rather intriguing analytic behaviors. For example, while the parameter λ held great significance in the proofs of Fredholm's theorems, it may be entirely insignificant in the discussion of nonlinear integral equations with a similar structure. Our examples will also show that nonlinear integral equations can have an infinite family of solutions.

Any discussion of nonlinear integral equations should include some of the techniques by which they are routinely solved. In the case that the integrand in Eq. (6.1) can be factored as

$$G(x,t,\phi(t)) = K(x,t)F(t,\phi(t)),$$

we show how to obtain solutions if $K(x,t)$ is either separable or symmetric and positive. Converting a nonlinear integral equation to a nonlinear differential equation is also fruitful.

Hopefully, the discussion that is presented here will serve to spur the interest and curiosity of the reader to continue the study of this particularly fascinating area of mathematics.

6.1 Tools of the Trade

In this chapter, the reader should be familiar with the following topics:

- *Metrics, metric spaces, and completeness*: A *metric* on a set X is a mapping $d: X \times X \to \mathbb{R}$ with the following properties:

  1. $d(x,y) \geq 0$ for all $x,y \in X$.
  2. $d(x,y) = 0$ if and only if $x = y$.
  3. $d(x,y) = d(y,x)$ for all $x,y \in X$.
  4. $d(x,z) \leq d(x,y) + d(y,z)$ for all $x,y,z \in X$.

  A metric is also called a *distance function* and $d(x,y)$ is called the *distance* between x and y.

  A *metric space*[1] is a set endowed with a metric. The elements of X are called the *points* of X.

  A sequence $\{x_n\}$ in a metric space X is called a *d-Cauchy sequence in X* if, for every $\varepsilon > 0$, there exists an integer N (which depends upon ε) such that $d(x_m, x_n) < \varepsilon$ whenever $n, m \geq N$.

  A metric d for a metric space X is called *complete* if every d-Cauchy sequence in X converges to an element of X. A metric space X is called *d-complete* if d is a complete metric for X.

  In particular, we note that the space $\mathcal{C}[a,b]$ of functions that are continuous on the closed interval $[a,b]$ is complete with respect to the metric

  $$d(x(t), y(t)) = \max_{a \leq t \leq b} |x(t) - y(t)|.$$

- *Contraction mappings*: Let X be a metric space with metric d, and let $\mathbf{A}$ be a mapping from X into itself. Then $\mathbf{A}$ is called a *contraction mapping* if

  $$d(\mathbf{A}x, \mathbf{A}y) \leq \alpha\, d(x,y)$$

  for all $x,y \in X$, where $0 < \alpha < 1$. The constant α does not depend in any way on the elements of X. The possibility that $\alpha = 0$ has been omitted here. For if $\alpha = 0$, then $d(\mathbf{A}x, \mathbf{A}y) = 0$ for all $x,y \in X$, i.e., $\mathbf{A}$ is a constant mapping. If X is a space consisting of functions, then $\mathbf{A}$ is called a *contraction operator*.
- *Continuous mappings*: Let X be a metric space with metric d, and let $\mathbf{A}$ be a mapping from X into itself. Then $\mathbf{A}$ is *continuous* at the point $x_0 \in X$ if, for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $d(\mathbf{A}x, \mathbf{A}y) < \varepsilon$ whenever $d(x,y) < \delta$. If S is a subset of X and $\mathbf{A}$ is continuous at every point of S, then we say that $\mathbf{A}$ is continuous on S.
- *Fixed points and fixed point theorems*: Let $\mathbf{A}$ be a mapping from a metric space X into itself. If $\mathbf{A}x_0 = x_0$, then x_0 is called a *fixed point* of the mapping $\mathbf{A}$.

[1] More precisely, a metric space is a topological space whose topology is induced by a metric.

It is natural to ask whether mappings have fixed points. Many theorems that provide necessary conditions on metric spaces and their self-mappings for the existence of fixed points have been established. Among them is the following result known as Banach's Fixed Point Theorem. It is also known as the Principle of Contraction.

Theorem 6.1.1 (Banach). *Let X be a complete metric space endowed with the metric d. If $\mathbf{A}$ is a contraction mapping from X into itself, then there exists a unique point $x_0 \in X$ such that $\mathbf{A}x_0 = x_0$.*

Proof. Let $x \in X$ be arbitrary. Generate a sequence in X by setting $x_1 = \mathbf{A}x$ and $x_{n+1} = \mathbf{A}x_n$ for $n \geq 1$. Since $\mathbf{A}$ is a contraction mapping, there exists a constant α $(0 < \alpha < 1)$ such that

$$d(x_1, x_2) = d(\mathbf{A}x, \mathbf{A}x_1) \leq \alpha\, d(x, x_1) = \alpha\, d(x, \mathbf{A}x).$$

A short inductive argument establishes the inequality

$$d(x_n, x_{n+1}) \leq \alpha^n\, d(x, \mathbf{A}x)$$

for all $n \geq 1$. Since every metric satisfies the triangle inequality, we also have

$$\begin{aligned} d(x_n, x_{n+p}) &\leq d(x_n, x_{n+1}) + \cdots + d(x_{n+p-1}, x_{n+p}) \\ &\leq \alpha^n\,(1 + \alpha + \cdots + \alpha^{p-1})\, d(x, \mathbf{A}x) \end{aligned}$$

for every integer $p \geq 1$. It follows that

$$d(x_n, x_{n+p}) \leq \frac{\alpha^n}{1-\alpha}\, d(x, \mathbf{A}x), \tag{6.3}$$

from which we deduce that $d(x_n, x_{n+p}) \to 0$ as $n \to +\infty$ for any given integer $p \geq 1$. Thus, the sequence $\{x_n\}_{n=1}^{\infty}$ is d-Cauchy in X. Since we have assumed that X is d-complete, there exists an element $x_0 \in X$ such that

$$x_0 = \lim_{n \to +\infty} x_n.$$

To show that $\mathbf{A}x_0 = x_0$, we show that $d(x_0, \mathbf{A}x_0) = 0$. Let $\varepsilon > 0$ be arbitrary. In view of the convergence of the sequence $\{x_n\}_{n=1}^{\infty}$, there exists a positive integer N such that $d(x_0, x_n) < \varepsilon/2$ and $d(x_0, x_{n-1}) < \varepsilon/2$ for all $n \geq N$. Consequently,

$$\begin{aligned} d(x_0, \mathbf{A}x_0) &\leq d(x_0, x_n) + d(x_n, \mathbf{A}x_0) \\ &= d(x_0, x_n) + d(\mathbf{A}x_{n-1}, \mathbf{A}x_0) \\ &\leq d(x_0, x_n) + \alpha\, d(x_{n-1}, x_0) \\ &< \frac{\varepsilon}{2} + \alpha\, \frac{\varepsilon}{2} \\ &< \varepsilon. \end{aligned}$$

But since ε is arbitrary, it must be the case that $d(x_0, \mathbf{A}x_0) = 0$.

It is easy to prove by contradiction that x_0 is unique. Suppose that $\mathbf{A}$ has two or more distinct fixed points, say x_0 and y_0. Then,

$$d(x_0, y_0) = d(\mathbf{A}x_0, \mathbf{A}y_0) \leq \alpha\, d(x_0, y_0).$$

If $x_0 \neq y_0$, then $d(x_0, y_0) > 0$, and we conclude that $\alpha \geq 1$, contrary to the hypothesis that $0 < \alpha < 1$. Hence, x_0 is the unique fixed point of the contraction mapping $\mathbf{A}$, and the proof is complete. □

An estimate on the speed of convergence of the sequence $\{x_n\}_{n=1}^{\infty}$ can be obtained from inequality (6.3). If we let $p \to +\infty$, then we obtain the inequality

$$d(x_n, x_0) \leq \frac{\alpha^n}{1-\alpha}\, d(x, \mathbf{A}x),$$

which shows that the speed of convergence depends upon the choice of metric, the magnitude of α, and the distance between x and its image $\mathbf{A}x$ in X.

Not every continuous mapping from X into itself must be a contraction mapping. However, it may happen that some power $\mathbf{A}^n$ of $\mathbf{A}$ might be a contraction mapping, even if $\mathbf{A}$ is not. In this case, another fixed point theorem can be enunciated.

Theorem 6.1.2. *Let* $\mathbf{A}$ *be a continuous mapping from the complete metric space* X *into itself. If* $\mathbf{A}^n$ *is a contraction mapping for some positive integer* n*, then there exists a unique point* $x_0 \in X$ *such that* $\mathbf{A}x_0 = x_0$.

- *Lipschitz conditions*: A real-valued function satisfies a *Lipschitz condition* on a set E if the inequality

$$|f(x) - f(y)| \leq C\,|x - y|$$

holds for all $x, y \in E$, where C is a constant that is independent of x and y. For example, a function that has a bounded first derivative on E satisfies a Lipschitz condition as a consequence of the Mean Value Theorem, which states that there exists a value c between x and y for which

$$f(x) - f(y) = f'(c)\,(x - y).$$

A function of several variables can satisfy a Lipschitz condition on a set with respect to one of the variables. In this chapter, we will encounter functions $G(x,t,z)$ of three variables that satisfy a Lipschitz condition of the form

$$|G(x,t,z_1) - G(x,t,z_2)| \leq C\,|z_1 - z_2|$$

where $G(x,t,z)$ is defined on $Q(a,b) \times \mathbb{R}$. The constant C is independent of x, t, z_1, and z_2.

6.2 Existence and Uniqueness Theorems

The fixed point theorems that were stated in Sect. 6.1 can be invoked to establish existence and uniqueness theorems for nonlinear integral equations. In this section, we illustrate the method by which this can be accomplished.

6.2.1 Nonlinear Fredholm Integral Equations

Banach's Fixed Point Theorem can be used to prove the following result.

Theorem 6.2.1. *Suppose that $G(x,t,z)$ is defined and continuous on the set $Q(a,b)\times\mathbb{R}$ and that it satisfies a Lipschitz condition of the form*

$$|G(x,t,z_1)-G(x,t,z_2)|<C\,|z_1-z_2|\,.$$

Suppose further that $f\in\mathcal{C}[a,b]$. Then the nonlinear Fredholm integral equation

$$\phi(x)=f(x)+\lambda\int_a^b G(x,t,\phi(t))\,\mathrm{d}t \tag{6.4}$$

has a unique solution on the interval $[a,b]$ whenever $|\lambda|<1/(C\,(b-a))$.

Proof. The space $\mathcal{C}[a,b]$ consisting of the real-valued continuous functions that are defined on the interval $[a,b]$ is a complete metric space with the metric

$$d(u,v)=\max_{a\le x\le b}|u(x)-v(x)|.$$

If it can be shown that the operator

$$\mathbf{F}\colon\,\mathbf{C}[a,b]\to\mathbf{C}[a,b]$$

defined by

$$\mathbf{F}u=f(x)+\lambda\int_a^b G(x,t,u(t))\,\mathrm{d}t$$

is a contraction operator for the restricted values of λ prescribed in the statement of the theorem, then it will follow immediately as an application of Banach's Fixed Point Theorem (Theorem 6.1.1) that $\mathbf{F}$ has a unique fixed point, i.e., that the integral equation in the statement of the theorem has a unique solution. Observe that for any $u,v\in\mathcal{C}[a,b]$ and any $x\in[a,b]$, we have

$$\begin{aligned}|\mathbf{F}u(x)-\mathbf{F}v(x)|&=|\lambda|\left|\int_a^b\left(G(x,t,u(t))-G(x,t,v(t))\right)\mathrm{d}t\right|\\&\le|\lambda|\,C\,(b-a)\,d(u,v).\end{aligned}$$

It follows directly from this inequality that

$$d(\mathbf{F}u, \mathbf{F}v) \leq |\lambda| C (b-a) d(u,v).$$

Hence, if $\alpha = |\lambda| C (b-a) < 1$, then $\mathbf{F}$ is a contraction operator from $\mathcal{C}[a,b]$ into itself, and the proof is complete. □

In the proof of the Theorem 6.2.1, a unique fixed point was determined as the limit of a sequence $\{x_n\}_{n=1}^{\infty}$ of points generated by the operator $\mathbf{A}$. Analogously, the unique solution of the nonlinear integral equation (6.4) is determined as the limit of a sequence of continuous functions generated by the operator $\mathbf{F}$. Choose $\phi_0(x)$. If we set

$$\phi_1(x) = \mathbf{F}\phi_0(x) = f(x) + \lambda \int_a^b G(x,t,\phi_0(t))\,\mathrm{d}t$$

and

$$\phi_{n+1}(x) = \mathbf{F}\phi_n(x)$$

for all $n \geq 1$, then the sequence $\{\phi_n(x)\}_{n=0}^{\infty}$ in $\mathcal{C}[a,b]$ converges uniformly to a limit function $\phi(x)$ which will be the solution to the integral equation. In essence, the Method of Successive Substitution is valid under the hypotheses of Theorem 6.2.1.

6.2.2 *Nonlinear Volterra Integral Equations*

Fixed point theorems can also be used to prove the existence and uniqueness of solutions to nonlinear Volterra integral equations. In the theorem below, note especially that there are no restrictions on the value of λ.

Theorem 6.2.2. *Suppose that $G(x,t,z)$ is defined and continuous on the set $T(a,b) \times \mathbb{R}$ and that it satisfies a Lipschitz condition of the form*

$$|G(x,t,z_1) - G(x,t,z_2)| < C\,|z_1 - z_2|\,.$$

Suppose further that $f \in \mathcal{C}[a,b]$. Then the nonlinear Volterra integral equation

$$\phi(x) = f(x) + \lambda \int_a^x G(x,t,\phi(t))\,\mathrm{d}t$$

has a unique solution on the interval $[a,b]$ for every value of λ, where $a \leq x \leq b$.

Proof. The space $\mathcal{C}[a,b]$ consisting of the real-valued continuous functions that are defined on the interval $[a,b]$ is a complete metric space with the metric

$$d(u,v) = \max_{a \leq x \leq b} |u(x) - v(x)|.$$

If it can be shown that a suitable power of the continuous operator

$$\mathbf{V}\colon\ \mathbf{C}[a,b] \to \mathbf{C}[a,b]$$

defined by

$$\mathbf{V}u = f(x) + \lambda \int_a^x G(x,t,u(t))\,\mathrm{d}t$$

is a contraction operator for any value of λ, then it will follow immediately as an application of Theorem 6.1.2 that $\mathbf{V}$ has a unique fixed point, i.e., that the integral equation in the statement of the theorem has a unique solution. Observe that for any $u, v \in \mathcal{C}[a,b]$ and any $x \in [a,b]$, we have

$$\begin{aligned} |\mathbf{V}u(x) - \mathbf{V}v(x)| &= |\lambda| \left| \int_a^x (G(x,t,u(t)) - G(x,t,v(t)))\,\mathrm{d}t \right| \\ &\le |\lambda| C\,(x-a)\,d(u,v). \end{aligned}$$

After again invoking the Lipschitz condition and then integrating the result, we obtain

$$\begin{aligned} |\mathbf{V}^2u(x) - \mathbf{V}^2v(x)| &= |\lambda| \left| \int_a^x (G(x,t,\mathbf{V}u(t)) - G(x,t,\mathbf{V}v(t)))\,\mathrm{d}t \right| \\ &\le |\lambda| C \int_a^x |\mathbf{V}u(t) - \mathbf{V}v(t)|\,\mathrm{d}t \\ &\le \frac{|\lambda|^2 C^2 (x-a)^2}{2!}\,d(u,v). \end{aligned}$$

In general, after n such integrations, we obtain

$$|\mathbf{V}^n u(x) - \mathbf{V}^n v(x)| \le \frac{|\lambda|^n C^n (x-a)^n}{n!}\,d(u,v),$$

from which it follows that

$$d(\mathbf{V}^n u, \mathbf{V}^n v) \le \frac{|\lambda|^n C^n (b-a)^n}{n!}\,d(u,v).$$

Hence, if n is sufficiently large, then $\mathbf{V}^n$ is a contraction operator from $\mathcal{C}[a,b]$ into itself, and the proof is complete. □

6.2.3 Hammerstein's Equation

In Sect. 5.2.1, it was shown that the boundary value problem that consists of the ordinary differential equation

$$\phi''(x) + F[x, \phi(x)] = 0$$

together with the boundary conditions $\phi(0) = \phi(1) = 0$ can be converted into the integral equation

$$\phi(x) = \int_0^1 T(x,t)\, F[t, \phi(t)]\, \mathrm{d}t,$$

where $T(x,t)$ is the continuous, symmetric, positive, "triangular" kernel given by

$$T(x,t) = \begin{cases} t(1-x) & \text{if } 0 \le t \le x \le 1 \\ x(1-t) & \text{if } 0 \le x \le t \le 1. \end{cases}$$

Hammerstein investigated a generalization of this converted equation by considering integral equations in the standardized form

$$\phi(x) = \int_0^1 K(x,t)\, F[t, \phi(t)]\, \mathrm{d}t, \tag{6.5}$$

where $K(x,t)$ is a symmetric, positive kernel, and reasonable restrictions are imposed on $F[t,z]$.

It is easy to choose functions $F[t,z]$ for which a solution need not exist. It is also evident that if $F[t,0] \equiv 0$, then $\phi(x) \equiv 0$ is a trivial solution to the equation. It is therefore natural to ask for reasonable hypotheses under which an equation of this type has nontrivial solutions. A comprehensive discussion of existence and uniqueness questions is beyond the scope of this textbook.[2] However, it is nonetheless possible to state one of several versions of a typical existence theorem and then to give an indication of the arguments involved in proving it, since Hammerstein's assumption that $K(x,t)$ is symmetric allows for an artful application of the Hilbert–Schmidt theory that was introduced in Chap. 3.

Theorem 6.2.3 (Hammerstein). *Suppose that $K(x,t)$ is a continuous, positive, symmetric kernel. If $F(t,z)$ is continuous and satisfies the inequality*

$$|F[t,z]| \le C_1\, |z| + C_2,$$

[2]For a treatment of these issues, see Tricomi, *Integral Equations*, pp. 202–213.

where C_1 and C_2 are positive and C_1 is smaller than the smallest eigenvalue of $K(x,t)$, then the nonlinear integral equation (6.5) has at least one solution. Furthermore, if either

(a) The function $F[t,z]$ is a nondecreasing function of z for any fixed $t \in (0,1)$ or

(b) The function $F[t,z]$ satisfies uniformly a Lipschitz condition of the form

$$|F[t,z_1] - F[t,z_2]| < C\,|z_1 - z_2| \quad (0 < C < \lambda_1),$$

then the nonlinear integral equation (6.5) has at most one solution.

If we assume that $F[x,\phi(x)] \in \mathcal{R}^2[0,1]$, then an invocation of the Hilbert–Schmidt Theorem (Theorem 3.4.1) allows us to conclude that the solution $\phi(x)$ has the representation

$$\phi(x) = \sum_{n=1}^{\infty} c_n\,\phi_n(x),$$

where $\{\phi_n(x)\}_{n=1}^{\infty}$ is a mutually orthonormal system of eigenfunctions that must correspond to the ordered positive eigenvalues $\lambda_1, \lambda_2, \ldots$ of the kernel $K(x,t)$.

A formula for each of the coefficients c_n can be given in terms of them all. Recall that, for each $n \geq 1$, we have

$$c_n = \int_0^1 \phi(x)\,\phi_n(x)\,\mathrm{d}x$$

due to the orthonormality of the eigenfunctions. If we substitute the right-hand side of Eq. (6.5) into this representation, we obtain

$$\begin{aligned} c_n &= \int_0^1 \left(\int_0^1 K(x,t)\,F[t,\phi(t)]\,\mathrm{d}t \right) \phi_n(x)\,\mathrm{d}x \\ &= \int_0^1 F[t,\phi(t)] \left(\int_0^1 K(x,t)\,\phi_n(x)\,\mathrm{d}x \right) \mathrm{d}t \\ &= \frac{1}{\lambda_n} \int_0^1 F[t,\phi(t)]\,\phi_n(t)\,\mathrm{d}t \\ &= \frac{1}{\lambda_n} \int_0^1 F\left[t, \sum_{m=1}^{\infty} c_m\,\phi_m(t) \right] \phi_n(t)\,\mathrm{d}t. \end{aligned}$$

If the kernel $K(x,t)$ has a finite bilinear expansion (see Sect. 3.3) consisting of N terms, then these representations for the coefficients $c_1,\ldots,c_N$ constitute a finite nonlinear system of N equations in N unknowns. However, if $K(x,t)$ has an infinite number of eigenvalues, then these representations constitute a nonlinear system of an infinite number of equations in an infinite number of unknowns. Each solution to the system, whether finite or infinite, should give rise to a solution to the Hammerstein equation.

The initial step in the process of establishing the existence of a solution is to truncate the infinite system of equations. For every positive integer p, we consider the finite system

$$c_{p,n} = \frac{1}{\lambda_n} \int_0^1 F\left[t, \sum_{m=1}^{p} c_{p,m}\,\phi_m(t)\right] \phi_n(t)\,\mathrm{d}t \quad (1 \le n \le p).$$

By means of an elegant minimization argument, Hammerstein showed that this finite system always has at least one solution, $(c_{p,1},\ldots,c_{p,p})$. As a consequence, a sequence of functions of the form

$$\psi_p(x) = \sum_{n=1}^{p} c_{p,n}\,\phi_n(x) \quad (1 \le p < +\infty)$$

can be constructed. This sequence has several valuable properties. For example, it can be shown that there exists a constant B such that $\|\psi_p\|_2 \le B$ for all $p \ge 1$, i.e., that the sequence $\{\psi_p\}_{p=1}^{\infty}$ is bounded. Although it can also be shown that

$$\lim_{p\to+\infty} \left(\psi_p(x) - \int_0^1 K(x,t)\,F[t,\psi_p(t)]\,\mathrm{d}t \right) = 0$$

uniformly on the interval $[0,1]$, the sequence $\{\psi_p(x)\}$ itself may not converge. However, it can be shown (via an argument involving equicontinuity and an application of Ascoli's Theorem) there will always exist a subsequence of it that converges uniformly on the interval $[0,1]$ to a continuous limit function $\psi(x)$, and this fact leads to the proof of the theorem.

6.3 Illustrative Examples

In this section, we consider a large variety of interesting examples in order to illustrate many of the new phenomena and intriguing analytic behaviors that can and do occur in the investigation of nonlinear integral equations. Hopefully, the discussion of these examples will spur the interest and curiosity of the reader to continue the study of this particularly fascinating area of mathematics.

- *Example 1*: Many of the important nonlinear integral equations assume the general form

$$L(x,\phi(x)) = \int_a^b G(x,t,\phi(t))\,\mathrm{d}t,$$

where $L(x,z)$ and $G(x,t,z)$ are nonlinear functions of z. In some cases, it is advantageous to incorporate the nonlinearity on the left-hand side of the equation into the integrand. If it is possible to invert the relationship $\psi(x) = L(x,\phi(x))$ to obtain $\phi(x) = J(x,\psi(x))$, then the integral equation can be reformulated as

$$\psi(x) = \int_a^b G(x,t,J(t,\psi(t)))\,\mathrm{d}t.$$

- *Example 2*: If the integrand $G(x,t,\phi(t))$ in the nonlinear equation (6.1) can be factored as

$$G(x,t,\phi(t)) = K(x,t)\,F(t,\phi(t)),$$

where $K(x,t)$ is separable, then a solution can often be found directly. For example, if $K(x,t) = a(x)\,b(t)$, then the solution must have the form $\phi(x) = f(x) + \lambda\, c\, a(x)$, where

$$c = \int_a^b b(t)\,F(t,\phi(t))\,\mathrm{d}t;$$

substituting $\phi(t)$ into this integrand yields a necessary condition, usually nonlinear in nature, that must be satisfied by the value c.

Consider the nonlinear Fredholm integral equation

$$\phi(x) = 1 + 2x + \lambda \int_0^1 xt\left(1+\phi^2(t)\right)\,\mathrm{d}t$$

where we assume that λ is a real parameter.

Every solution to this equation must have the form $\phi(x;\lambda) = 1 + (2+\lambda c)x$, where we have set

$$c = \int_0^1 t\,\left(1+\phi^2(t)\right)\,\mathrm{d}t.$$

Given the form of the solution, this integral condition is equivalent to the quadratic equation

$$3\lambda^2 c^2 + (20\lambda - 12)\,c + 40 = 0,$$

whose roots in terms of the parameter λ are

$$c_1(\lambda) = \frac{2}{3}\left(\frac{3-5\lambda-\sqrt{9-30\lambda-5\lambda^2}}{\lambda^2}\right)$$

and

$$c_2(\lambda) = \frac{2}{3}\left(\frac{3-5\lambda+\sqrt{9-30\lambda-5\lambda^2}}{\lambda^2}\right).$$

Note that $c_1(\lambda)$ has a removable discontinuity at $\lambda = 0$, but $c_2(\lambda)$ does not. However, this presents no difficulty, since $\phi(x;0) = 1+2x$.

In order to obtain the rest of the real-valued solutions to the integral equation, we must require that $c_1(\lambda)$ and $c_2(\lambda)$ be real. The discriminant $\Delta(\lambda) = 9 - 30\lambda - 5\lambda^2$ vanishes at the values

$$\lambda_1 = -\frac{3}{5}\left(5+\sqrt{30}\right) = -6.2863\ldots$$

and

$$\lambda_2 = -\frac{3}{5}\left(5-\sqrt{30}\right) = +0.2863\ldots$$

and is positive between them. Thus,

$$c_1(\lambda_1) = c_2(\lambda_1) = \frac{10}{9}\left(6-\sqrt{30}\right) = +0.5808\ldots$$

so that

$$\phi(x;\lambda_1) = 1+\left(2-\frac{2}{3}\sqrt{30}\right)x,$$

and

$$c_1(\lambda_2) = c_2(\lambda_2) = \frac{10}{9}\left(6+\sqrt{30}\right) = +12.7525\ldots$$

so that

$$\phi(x;\lambda_2) = 1+\left(2+\frac{2}{3}\sqrt{30}\right)x.$$

The solutions $\phi(x;0)$, $\phi(x;\lambda_1)$, and $\phi(x;\lambda_2)$ are unique. However, if either $\lambda_1 < \lambda < 0$ or $0 < \lambda < \lambda_2$, then there are exactly two solutions, given by

$$\begin{aligned}\phi^-(x;\lambda) &= 1+\left(2+\frac{2}{3}\left(\frac{3-5\lambda-\sqrt{9-30\lambda-5\lambda^2}}{\lambda^2}\right)\lambda\right)x \\ &= 1+\left(\frac{6-4\lambda-2\sqrt{9-30\lambda-5\lambda^2}}{3\lambda}\right)x\end{aligned}$$

and

$$\phi^+(x;\lambda) = 1 + \left(2 + \frac{2}{3}\left(\frac{3 - 5\lambda + \sqrt{9 - 30\lambda - 5\lambda^2}}{\lambda^2}\right)\lambda\right)x$$
$$= 1 + \left(\frac{6 - 4\lambda + 2\sqrt{9 - 30\lambda - 5\lambda^2}}{3\lambda}\right)x.$$

If either $-\infty < \lambda < \lambda_1$ or $\lambda_2 < \lambda < +\infty$, then there are no real-valued solutions.

To summarize, we have shown that the solutions to the integral equation constitute a family of straight lines passing through the point $(0,1)$. Finally, we note that the only constant solution is $\phi^-(x;-2) = 1$.

- *Example 3*: The concept of an eigenvalue played an extremely significant role in the discussion of the Fredholm theorems earlier in this text. For example, the Fourth Fredholm Theorem, considered in Chap. 2, states that the set of eigenvalues of the homogeneous equation

$$\phi(x) = \lambda \int_a^b K(x,t)\,\phi(t)\,\mathrm{d}t$$

is at most countable and has no finite limit point.

By way of contrast, consider the nonlinear integral equation

$$\phi(x) = \lambda \int_a^b K(x,t)\,\phi^n(t)\,\mathrm{d}t \quad (n \neq 1)$$

If we set $\phi(x) = \mu\,\psi(x)$, where μ is chosen so that $\lambda\,\mu^{n-1} = 1$, then this nonlinear equation is readily transformed into the equation

$$\psi(x) = \int_a^b K(x,t)\,\psi^n(t)\,\mathrm{d}t,$$

for which both of the values λ and μ are entirely insignificant.

- *Example 4*: The method used to solve linear integral equations that had separable kernels can be adapted to solve nonlinear integral equations as well. However, a important difference arises. While the method that we used to solve linear integral equations led to a linear system of equations, the method here almost always leads to a nonlinear system of equations.

Consider the nonlinear integral equation

$$\phi(x) = \int_0^\pi \sin(x+t)\,\phi^2(t)\,\mathrm{d}t.$$

An obvious solution to this equation is the trivial solution $\phi(x) \equiv 0$. To determine nontrivial solutions, we expand the integrand to obtain

$$\phi(x) = \int_0^{\pi} (\sin x \cos t + \cos x \sin t)\, \phi^2(t)\, \mathrm{d}t,$$

from which we conclude that every possible solution must have the form

$$\phi(x) = c_1 \sin x + c_2 \cos x,$$

where

$$c_1 = \int_0^{\pi} \cos t\, \phi^2(t)\, \mathrm{d}t \quad \text{and} \quad c_2 = \int_0^{\pi} \sin t\, \phi^2(t)\, \mathrm{d}t.$$

After substituting $\phi(t)$ into these necessary conditions and simplifying the results, we find that c_1 and c_2 must simultaneously satisfy the nonlinear equations

$$c_1 \left(1 - \frac{4}{3} c_2\right) = 0 \quad \text{and} \quad 2c_1^2 + \left(c_2 - \frac{3}{4}\right)^2 = \frac{9}{16}.$$

In the c_1c_2-plane, the first equation represents two straight lines, and the second equation represents an ellipse centered at the point $\left(0, \frac{3}{4}\right)$. These two equations have four solutions (c_1, c_2):

$$(0,0), \qquad \left(0, \frac{3}{2}\right), \qquad \left(+\frac{3\sqrt{2}}{8}, \frac{3}{4}\right), \quad \text{and} \quad \left(-\frac{3\sqrt{2}}{8}, \frac{3}{4}\right).$$

Corresponding to $c_1 = 0$, we have the two solutions

$$\phi(x) = 0 \quad \text{and} \quad \phi(x) = \frac{3}{2} \cos x;$$

corresponding to $c_2 = \frac{3}{4}$, we have the other two solutions

$$\phi(x) = +\frac{3\sqrt{2}}{8} \sin x + \frac{3}{4} \cos x \quad \text{and} \quad \phi(x) = -\frac{3\sqrt{2}}{8} \sin x + \frac{3}{4} \cos x.$$

- *Example 5*: Nonlinear integral equations can have an *infinite* family of solutions. Consider the nonlinear integral equation

$$\phi(x) = \int_0^{\pi} \cos(x - t)\, \phi^3(t)\, \mathrm{d}t.$$

An obvious solution to this equation is the trivial solution $\phi(x) \equiv 0$. To determine nontrivial solutions, we expand the integrand to obtain

$$\phi(x) = \int_0^\pi (\cos x \cos t + \sin x \sin t)\, \phi^3(t)\, \mathrm{d}t,$$

from which we conclude that every possible solution must have the form

$$\phi(x) = c_1 \cos x + c_2 \sin x,$$

where

$$c_1 = \int_0^\pi \cos t\, \phi^3(t)\, \mathrm{d}t \quad \text{and} \quad c_2 = \int_0^\pi \sin t\, \phi^3(t)\, \mathrm{d}t.$$

After substituting $\phi(t)$ into these necessary conditions and simplifying the results, we find that c_1 and c_2 must simultaneously satisfy the nonlinear conditions

$$c_1 \left(1 - \frac{3\pi}{8}\left(c_1^2 + c_2^2\right)\right) = 0 \quad \text{and} \quad c_2 \left(1 - \frac{3\pi}{8}\left(c_1^2 + c_2^2\right)\right) = 0.$$

If $c_1 \neq 0$ and $c_2 = 0$, then the first condition implies that

$$\phi(x) = \pm 2\sqrt{\frac{2}{3\pi}} \cos x.$$

If $c_1 = 0$ and $c_2 \neq 0$, then the second condition implies that

$$\phi(x) = \pm 2\sqrt{\frac{2}{3\pi}} \sin x.$$

If $c_1 \neq 0$ and $c_2 \neq 0$, then the point (c_1, c_2) in the c_1c_2-plane must belong to the circle $c_1^2 + c_2^2 = \frac{8}{3\pi}$. If we set

$$c_1 = 2\sqrt{\frac{2}{3\pi}} \cos \alpha \quad \text{and} \quad c_2 = 2\sqrt{\frac{2}{3\pi}} \sin \alpha,$$

where α is real and arbitrary, then

$$\phi(x) = 2\sqrt{\frac{2}{3\pi}} (\cos x \cos \alpha + \sin x \sin \alpha) = 2\sqrt{\frac{2}{3\pi}} \cos(x - \alpha).$$

This general solution includes the previous two solutions as special cases. Note that the integral equation in this example has an infinite number of solutions, in contrast to the integral equation in the previous example which had only four.

- *Example 6*: A simple example shows that nonlinear integral equations can have complex-valued solutions in addition to real-valued solutions.

  Consider the nonlinear integral equation

$$\phi(x) = \int_0^\pi \cos(x-t)\,\phi^4(t)\,\mathrm{d}t.$$

  An obvious solution to this equation is the trivial solution $\phi(x) \equiv 0$. To determine nontrivial solutions, we expand the integrand to obtain

$$\phi(x) = \int_0^\pi (\cos x \cos t + \sin x \sin t)\,\phi^4(t)\,\mathrm{d}t,$$

  from which we conclude that every possible solution must have the form

$$\phi(x) = c_1 \cos x + c_2 \sin x,$$

  where

$$c_1 = \int_0^\pi \cos t\,\phi^4(t)\,\mathrm{d}t \quad \text{and} \quad c_2 = \int_0^\pi \sin t\,\phi^4(t)\,\mathrm{d}t.$$

  After substituting $\phi(t)$ into these necessary conditions and simplifying the results, we find that c_1 and c_2 must simultaneously satisfy the nonlinear equations

$$c_1 = \frac{8}{15}\,c_1 c_2\left(3c_1^2 + 2c_2^2\right) \quad \text{and} \quad c_2 = \frac{2}{15}\left(3c_1^4 + 12c_1^2c_2^2 + 8c_2^4\right).$$

  The first equation implies that if $c_2 = 0$, then $c_1 = 0$, which leads to the trivial solution. Hence, we assume that $c_2 \neq 0$.

  If $c_1 = 0$, then the second equation implies that $c_2^3 = \frac{15}{16}$.

  If $c_1 \neq 0$, then the first equation can be rewritten as

$$c_1^2 = \frac{15 - 16c_2^3}{24\,c_2}.$$

  After substituting this expression for c_1^2 into the second equation, we find that c_2 must satisfy the relation

$$\frac{(15 - 16c_2^3)^2}{1440\,c_2^2} = 0,$$

  i.e., $c_2^3 = \frac{15}{16}$, implying that $c_1^2 = 0$, contrary to assumption.

We conclude that the only solutions to the pair of nonlinear equations are $(0,0)$ and $(0,\zeta\sqrt[3]{\frac{15}{16}})$, where $\zeta^3=1$. Therefore, the only nontrivial solutions to the integral equation are

$$\phi(x)=\zeta\sqrt[3]{\frac{15}{16}}\sin x.$$

- *Example 7*: In this example, we illustrate some of the difficulties that can arise when the nonlinearity occurs outside of the integrand.

  Suppose that we desire all real-valued solutions to the nonlinear integral equation

$$\phi^2(x)=1+x^2+\lambda\int_0^1 xt\,\phi(t)\,\mathrm{d}t.$$

If $\lambda=0$, then it is clear that there are exactly two real-valued solutions, $\phi(x)=\pm\sqrt{1+x^2}$. Furthermore, it can be observed that, regardless of the value of λ, every nontrivial real-valued solution must satisfy the relation

$$\phi^2(x)=1+\lambda cx+x^2,$$

where we have set

$$c=\int_0^1 t\,\phi(t)\,\mathrm{d}t.$$

Real-valued solutions defined on the interval $[0,1]$ can only exist if $p(x)=1+\lambda cx+x^2\geq 0$ there. On the one hand, if $\lambda c\geq -2$, then

$$p(x)=1+\lambda cx+x^2\geq 1-2x+x^2=(1-x)^2\geq 0.$$

On the other hand, if $\lambda c<-2$ and the zeroes of $p(x)$ are r_1 and r_2, then $r_1+r_2=-\lambda c>+2$ and $r_1r_2=1$, implying that $p(x)$ has a simple zero in the interior of the interval $(0,1)$. Thus, real-valued solutions can only exist on the interval $[0,1]$ if $\lambda c\geq -2$.

Define

$$\phi^+(x)=+\sqrt{1+\lambda cx+x^2}\quad\text{and}\quad\phi^-(x)=-\sqrt{1+\lambda cx+x^2}.$$

Set

$$c^+=\int_0^1 t\,\phi^+(t)\,\mathrm{d}t=\int_0^1 t\sqrt{1+\lambda c^+t+t^2}\,\mathrm{d}t,$$

and observe that

$$c^+ = \int_0^1 t\sqrt{1+\lambda c^+ t+t^2}\,\mathrm{d}t \geq \int_0^1 t\sqrt{1-2t+t^2}\,\mathrm{d}t = \frac{1}{6}.$$

Note especially that if $\lambda c^+ = -2$, then $c^+ = \frac{1}{6}$ and that $\lambda = -12$.

With a similar definition for c^-, we also have $c^- \leq -\frac{1}{6}$. Also, if $\lambda c^- = -2$, then $c^- = -\frac{1}{6}$ and $\lambda = +12$.

We can now show that there is a unique value of c^+ for which $\phi^+(x)$ is a solution to the integral equation on the interval $[0,1]$. To do this, we shall show that there is a unique solution c^+ to the equation

$$\frac{1}{\lambda} = \frac{\int_0^1 t\sqrt{1+\lambda c^+ t+t^2}\,\mathrm{d}t}{\lambda c^+}.$$

Define the auxiliary function

$$R(a) = \frac{\int_0^1 t\sqrt{1+at+t^2}\,\mathrm{d}t}{a} = \frac{1}{\sqrt{a}}\left(\int_0^1 t\sqrt{t+\frac{1+t^2}{a}}\,\mathrm{d}t\right).$$

It is clear that $R(a)$ is positive and strictly decreasing on the interval $0 < a < +\infty$, that $R(a) \to +\infty$ as $a \to 0+$, and that $R(a) \to 0$ as $a \to +\infty$. It is also clear that $R(a)$ is negative and strictly decreasing on the interval $-2 \leq a < 0$, that $R(a) \to -\frac{1}{12}$ as $a \to -2$, and that $R(a) \to -\infty$ as $a \to 0^-$. Given this information, we conclude that $R(a)$ has a uniquely defined inverse on its domain.

Given the range of $R(a)$, the equation $\frac{1}{\lambda} = R(a)$ has a uniquely defined solution $a = R^{-1}(\frac{1}{\lambda})$ whenever $-12 \leq \lambda < 0$ or $0 < \lambda < +\infty$. If we set $a = \lambda c^+$, then $c^+ = \frac{1}{\lambda}R^{-1}(\frac{1}{\lambda})$. We can now conclude that $\phi^+(x)$ is a solution to the integral equation for each value of λ in the given intervals.

In a similar manner, it can be shown that the equation

$$\frac{1}{\lambda} = \frac{-\int_0^1 t\sqrt{1+\lambda c^- t+t^2}\,\mathrm{d}t}{\lambda c^-}$$

has a unique solution c^- or, equivalently, that the equation $-\frac{1}{\lambda} = R(a)$ has a uniquely defined solution $a = R^{-1}(-\frac{1}{\lambda})$ whenever $-\infty < \lambda < 0$ or $0 < \lambda \leq 12$. If we set $a = \lambda c^-$, then $c^- = \frac{1}{\lambda}R^{-1}(-\frac{1}{\lambda})$. We can now also conclude that $\phi^-(x)$ is a solution to the integral equation for each value of λ in the given intervals.

Thus, the integral equation has exactly two solutions, $\phi^+(x)$ and $\phi^-(x)$, if $-12 \leq \lambda \leq +12$, exactly one solution $\phi^+(x)$ if $12 < \lambda < +\infty$, and exactly one solution $\phi^-(x)$ if $-\infty < \lambda < -12$.

- *Example 8*: We have seen in several examples above that determining the solutions to a nonlinear integral equation may involve determining the solutions to a nonlinear system of algebraic equations. Although it was a relatively painless process in the preceding examples, it may in general be quite difficult or even impossible to do so.

  Suppose that we seek all real-valued solutions to the nonlinear integral equation

$$\phi(x) = 1 + x + x^2 + \lambda \int_0^1 \left(1 + xt + x^2t^2\right) \phi^2(t)\,\mathrm{d}t.$$

  If $\lambda = 0$, then the unique solution is obviously $\phi(x) = 1 + x + x^2$. If $\lambda \neq 0$, then every solution is a quadratic polynomial of the form

$$\phi(x) = (1 + \lambda a) + (1 + \lambda b)x + (1 + \lambda c)x^2,$$

  where a, b, and c are positive constants satisfying the relations

$$a = \int_0^1 \phi^2(t)\,\mathrm{d}t, \qquad b = \int_0^1 t\,\phi^2(t)\,\mathrm{d}t, \quad \text{and} \quad c = \int_0^1 t^2\,\phi^2(t)\,\mathrm{d}t.$$

  After substituting the required form for $\phi(t)$, we find that these relations are equivalent to the coefficient conditions

$$\begin{aligned} 30a &= 111 + (110a + 65b + 47c)\lambda \\ &\quad + \left(30a^2 + 30ab + 10b^2 + 20ac + 15bc + 6c^2\right)\lambda^2, \\ 60b &= 149 + (130a + 94b + 74c)\lambda \\ &\quad + \left(30a^2 + 40ab + 15b^2 + 30ac + 24bc + 10c^2\right)\lambda^2, \end{aligned}$$

  and

$$\begin{aligned} 210c &= 401 + (329a + 259b + 214c)\lambda \\ &\quad + \left(70a^2 + 105ab + 42b^2 + 84ac + 70bc + 30c^2\right)\lambda^2. \end{aligned}$$

  As λ varies, each of these conditions represents a family of ellipsoids (real or imaginary) in abc-space. Since we seek only real-valued solutions to the integral equation, we assume that all three of these ellipsoids are real.

  Recall that the intersection of three real ellipsoids may consist of as many as eight points or possibly a closed curve. After conducting a standard numerical investigation with *Mathematica*, we find there are two distinct real solutions and six complex solutions if $-72.2074 \approx \lambda_{\min} < \lambda < 0$ or $0 < \lambda < \lambda_{\max} \approx 0.0854635615$. If $\lambda = \lambda_{\min}$ or $\lambda = \lambda_{\max}$, then there is only one solution. If $\lambda < \lambda_{\min}$ or $\lambda > \lambda_{\max}$, then all eight solutions are complex.

For example, if $\lambda = 0.01$, then the two real solutions to the nonlinear algebraic system are

$$(a, b, c) \approx (3.93694, 2.63726, 2.02611),$$

which leads to the first solution

$$\phi(x) \approx 1.03937 + 1.02637x + 1.02026x^2,$$

and

$$(a, b, c) \approx (4195.41, 2609.03, 1936.74),$$

which leads to the second solution

$$\phi(x) \approx 42.9541 + 27.0903x + 20.3674x^2.$$

Note that the first solution is close to the solution $\phi(x) = 1 + x + x^2$ which corresponds to $\lambda = 0$, but the second solution is not.

If $\lambda = \lambda_{\max}$, then the only solution is

$$\phi(x) \approx 2.28352 + 1.83160x + 1.62922x^2.$$

If $\lambda = \lambda_{\min}$, then the only solution is

$$\phi(x) \approx -0.0881 + 0.1287x + 0.2425x^2.$$

Another approach to solving this integral equation involves the use of the Method of Successive Approximation.

Let $\lambda = 0.01$, and choose $\phi_0(x) = 1 + x + x^2$. By direct computation, we find that

$$\phi_1(x) = 1 + x + x^2 + 0.01 \int_0^1 \left(1 + xt + x^2t^2\right) \phi_0(t)\, \mathrm{d}t$$

$$\approx 1.03700 + 1.02483x + 1.01909x^2.$$

The next two approximants are

$$\phi_2(x) \approx 1.03922 + 1.02628x + 1.02019x^2$$

and

$$\phi_3(x) \approx 1.03936 + 1.02637x + 1.02026x^2.$$

The third approximant compares quite favorably to the solution that was produced directly by solving the nonlinear algebraic system above. The student should be cautioned, however, that serious questions concerning the convergence of such sequences must be answered before employing this method routinely for other values of λ or for other integral equations.

- *Example 9*: It is well known that Maclaurin series can be used to solve ordinary differential equations. In this example, we show that they can also be used to solve nonlinear integral equations of Volterra type.

  Consider the integral equation

$$\phi^2(x) = \int_0^x (x-t)\,\phi(t)\,\mathrm{d}t.$$

  It is obvious that $\phi(x) \equiv 0$ is a solution to this equation. Since $\phi(0) = 0$, it is reasonable to seek a nontrivial solution of the form

$$\phi(x) = x^\alpha \left(\sum_{n=0}^{\infty} a_n x^n \right) = \sum_{n=0}^{\infty} a_n x^{n+\alpha},$$

  where $\alpha > 0$ and $a_0 \neq 0$, assuming that the series has a positive radius of convergence.

  On the one hand, we have

$$\begin{aligned}\int_0^x (x-t)\phi(t)\,\mathrm{d}t &= \sum_{n=0}^{\infty} \frac{1}{(n+\alpha+1)(n+\alpha+2)}\, a_n x^{n+\alpha+2} \\ &= \frac{1}{(\alpha+1)(\alpha+2)}\, a_0 x^{\alpha+2} + \frac{1}{(\alpha+2)(\alpha+3)}\, a_1 x^{\alpha+3} + \cdots\end{aligned}$$

  On the other hand, we also have

$$\begin{aligned}\phi^2(x) &= x^{2\alpha}\left(a_0^2 + 2a_0a_1x + \cdots + \sum_{k=0}^{n} a_k a_{n-k} x^n + \cdots \right) \\ &= a_0^2 x^{2\alpha} + 2a_0a_1 x^{2\alpha+1} + \left(2a_0a_2 + a_1^2\right) x^{2\alpha+2} + \cdots.\end{aligned}$$

  By comparing the first terms in these two series, we deduce that $\alpha = 2$ and that $a_0 = \frac{1}{12}$. If we use this information when comparing the second terms, we find that $a_1 = 0$. A short induction argument shows that $a_n = 0$ for all $n \geq 1$. Given our assumption about the form of the solution to the integral equation, it now follows that

$$\phi(x) = \frac{1}{12} x^2,$$

  as is easily checked.

- *Example 10*: In this example, we compare and contrast the Method of Successive Approximation with the direct substitution of an infinite series.

  Consider the nonlinear integral equation of Volterra type

$$\phi(x) = x + \int_0^x (x-t)^2\,\phi^2(t)\,\mathrm{d}t.$$

Although this equation is a bit more complicated than the equation in the previous example, due to the inhomogeneous term and the nonlinearity in the integrand, we will find that both of the employed methods can still produce some very useful information. Since $\phi(0) = 0$, a reasonable initial approximant to the solution $\phi(x)$ is $\phi_0(x) = x$. Subsequent approximants can be determined from the recurrence formula

$$\phi_{n+1}(x) = x + \int_0^x (x-t)^2 \, \phi_n^2(t) \, \mathrm{d}t.$$

After a few simple integrations, we find that

$$\phi_1(x) = x + \frac{1}{30} x^5,$$

$$\phi_2(x) = x + \frac{1}{30} x^5 + \frac{1}{3780} x^9 + \frac{1}{772,200} x^{13},$$

and

$$\phi_3(x) = x + \frac{1}{30} x^5 + \frac{1}{3780} x^9 + \frac{31}{16,216,200} x^{13} + \cdots.$$

We can be reasonably sure that the coefficients of x^5 and x^9 in these approximants are the same as the initial coefficients in the representation for $\phi(x)$, since they do not change upon iteration.

Noticing that successive exponents in these approximants differ by four, it is reasonable to propose that a nontrivial solution to the integral equation has a series representation of the special form

$$\phi(x) = x \left(\sum_{n=0}^{\infty} a_n x^{4n} \right),$$

with a positive radius of convergence. After directly substituting this series into the integral equation and then equating its coefficients, we find that $a_0 = 1$ and that all of the subsequent coefficients satisfy the recursion relation

$$a_{n+1} = \frac{2}{(4n+3)(4n+4)(4n+5)} \left(\sum_{k=0}^{n} a_k \, a_{n-k} \right).$$

After some elementary calculations, we find that the first few coefficients in the series solution to the integral equation are

$$a_1 = \frac{1}{30}, \qquad a_2 = \frac{1}{3780}, \quad \text{and} \quad a_3 = \frac{31}{16,216,200}.$$

Note that these coefficients agree with those in $\phi_3(x)$. Note also that the direct substitution of an infinite series confirms the conjectured form of the solution. The recursion relationship that it provides allows us to compute as many coefficients as desired, with less computational effort and more accuracy than can be produced by using successive approximation.

- *Example 11*: If a parameter is present in a nonlinear integral equation of Volterra type, then the domain of the solution may depend upon that parameter. Other characteristics of the solution, such as the location of inflection points and the location of asymptotes, may depend upon the parameter as well.

  Consider the nonlinear integral equation

$$\phi(x) = \frac{1}{2}x^2 + \lambda \int_0^x t\, \mathrm{e}^{-\phi(t)}\, \mathrm{d}t.$$

  Regardless of the value of the parameter λ, we always have $\phi(0) = 0$. If $\phi(x) \equiv 0$ for all x, then necessarily $\lambda = -1$, by direct substitution.

  After differentiating the integral equation, we obtain

$$\phi'(x) = x\left(1 + \lambda\, \mathrm{e}^{-\phi(x)}\right),$$

  from which we deduce that $\phi'(0) = 0$, again regardless of the value of λ. After rewriting this equation in the form

$$\frac{\mathrm{e}^{\phi(x)}\, \phi'(x)}{\mathrm{e}^{\phi(x)} + \lambda} = x,$$

  we observe that if $\lambda = -1$, then the left-hand side of this equation is undefined at $x = 0$. If $\lambda \neq -1$, then the differential equation can be directly integrated to obtain the implicit solution

$$\ln\left|\frac{\mathrm{e}^{\phi(x)} + \lambda}{1 + \lambda}\right| = \frac{1}{2}x^2,$$

  which can in turn be easily solved to obtain the explicit solution

$$\phi(x;\lambda) = \ln\left[(1+\lambda)\,\mathrm{e}^{x^2/2} - \lambda\right].$$

  Note first that if $\lambda = -1$, then $\phi(x;-1) \equiv 0$. If the solution is rewritten in the form

$$\phi(x;\lambda) = \ln\left[1 + (1+\lambda)\left(\mathrm{e}^{x^2/2} - 1\right)\right],$$

  then, for each fixed value of x, we have $\phi(x;\lambda) \to 0$ as $\lambda \to -1$, due to the continuous dependence upon λ.

We turn to the dependence of the interval of validity of these solutions upon the value of the parameter λ. It is clear that $\phi(x;\lambda)$ is defined for all values of x if $\lambda = 0$ or $\lambda = -1$. This is also true if $0 < \lambda < +\infty$ or $-1 < \lambda < 0$, although for different technical reasons. However, if $-\infty < \lambda < -1$, then $\phi(x;\lambda)$ is only defined on the interval $-\tilde{\lambda} < x < +\tilde{\lambda}$, where

$$\tilde{\lambda} = \sqrt{2\ln\left(\frac{\lambda}{1+\lambda}\right)}.$$

The graph of the solution $\phi(x;\lambda)$ will have a vertical asymptote if either $x \to +\tilde{\lambda}^-$ or $x \to -\tilde{\lambda}^+$.

Other characteristics of the solution can be observed. For example, all solutions of the integral equation are even functions of x, a conclusion that was not at all clear from the form of the integral equation.

- *Example 12*: Consider the nonlinear integral equation

$$\phi(x) = f(x) + \lambda \int_a^b k(x,t,\phi(t))\,\mathrm{d}t,$$

where

$$k(x,t,z) = \sum_{i=1}^{n} a_i(x)\,b_i(t,z).$$

Then every solution to the integral equation must have the form

$$\phi(x) = f(x) + \lambda \sum_{i=1}^{n} c_i\,a_i(x),$$

where, for each $i = 1,\ldots,n$, we have

$$c_i = \int_a^b b_i\left(t, f(t) + \lambda \sum_{i=1}^{n} c_i\,a_i(t)\right)\mathrm{d}t.$$

Thus, to solve the integral equation, it is necessary to solve this nonlinear system of n equations in the n variables c_i. Some of the complications that can arise have been exhibited in the previous illustrative examples. Others will be discovered in the exercises to follow. It will be easy to compare and contrast the methods employed here with those that were introduced in Chap. 2 for linear integral equations with separable kernels.

- *Example 13*: A nonlinear integral equation of Volterra type of the first kind has the form

$$f(x) = \int_0^x G(x,t,\phi(t))\,\mathrm{d}t,$$

where $f(x)$ is a given continuous function on some interval $[0,b]$ with $f(0)=0$.

Formal differentiation of this integral equation yields

$$G(x,x,\phi(x)) = f'(x) - \int_0^x \frac{\partial G}{\partial x}(x,t,\phi(t))\,\mathrm{d}t,$$

an equation of the second kind which is not of the standard form considered in this chapter.

In order to show that the equation of the first kind has a unique solution, it is essential to show first that the corresponding integral equation of the second kind has a unique solution. This will be the case under the following familiar assumptions:

- $f(0)=0$.
- $f(x)$ and $f'(x)$ are continuous on the interval $[0,b]$.
- $G(x,t,z)$ and $\partial G(x,t,z)/\partial x$ are continuous on $T(0,b)\times\mathbb{R}$.
- $\partial G(x,t,z)/\partial x$ satisfies a Lipschitz condition with respect to z

together with two additional unfamiliar assumptions, namely:

- $G(x,x,z)$ is an invertible function of z for all $x\in[0,b]$.
- $|G(x,x,z_2)-G(x,x,z_1)|\le k\,|z_2-z_1|$ for some constant k, for each $x\in[0,b]$, and all z_1 and z_2.

Under all of these assumptions, the method of successive approximations can be applied to produce a sequence $\{\phi_n(x)\}_{n=0}^{\infty}$ that can be shown to converge to a unique solution due to an argument involving the use of a suitably constructed contraction mapping.

As a practical matter, the process of determining these approximating sequences may be quite complicated. After a zeroth approximant $\phi_0(x)$ is chosen, additional approximants $\phi_{n+1}(x)$ are then produced as solutions of the equation

$$G(x,x,\phi_{n+1}(x)) = f'(x) - \int_0^x \frac{\partial G}{\partial x}(x,t,\phi_n(t))\,\mathrm{d}t.$$

This can be done as a consequence of the invertibility assumption stated above. The assumptions on $G(x,t,z)$ are required to show that each $\phi_n(x)$ is continuous on $[0,b]$. While some of these assumptions are necessary as stated, others can often be adapted to special circumstances.

Section 6.3 Exercises

1. Consider the nonlinear integral equation

$$\phi(x) = f(x) + \int_0^x Q(x,t,\phi(t))\,\mathrm{d}t,$$

where

$$Q(x,t,z) = \sin(x-t)\,\frac{1}{1+z^3}.$$

(a) Show that $Q(x,t,z)$ satisfies a Lipschitz condition of the form

$$|Q(x,t,z_2) - Q(x,t,z_1)| \le C\,|z_2 - z_1|$$

on the interval $[0,\infty)$, where $C = \frac{2}{3}\sqrt[3]{2}$.

(b) Invoke Theorem 6.2.2 to conclude that the integral equation has a unique solution whenever $f(x)$ is continuous on $[0,\infty)$.

(c) Choose $f(x) = 1$. By using the differentiation tool stated in Sect. 4.1, differentiate the integral equation twice to show that $\phi(x)$ satisfies the nonlinear ordinary differential equation

$$\phi''(x) + \phi(x) = 1 + \frac{1}{1+\phi^3(x)},$$

together with the initial conditions $\phi(0) = 1$ and $\phi'(0) = 0$.

(d) Since $\phi''(0) = \frac{1}{2}$, we can choose $\phi_0(x) = 1 + \frac{1}{4}x^2$ to be an initial approximant, and then use it to compute the first approximant

$$\begin{aligned}\phi_1(x) &= 1 + \int_0^x \sin(x-t)\,\frac{1}{1+\phi_0^3(t)}\,\mathrm{d}t \\ &= 1 + \frac{1}{4}x^2 - \frac{7}{192}x^4 + \cdots.\end{aligned}$$

Note the inherent difficulties in performing the necessary integration. (It will be advantageous to use an appropriate partial sum of the Taylor series for $\sin(x-t)$ in doing so.)

(e) Show that $\phi'''(0) = 0$ and that $\phi''''(0) = -\frac{7}{8}$ by differentiating the differential equation in part (c) twice. Conclude that

$$\phi(x) = 1 + \frac{1}{4}x^2 - \frac{7}{192}x^4 + \cdots.$$

(f) Suppose that you needed to compute the coefficient of x^6. Would you prefer to compute $\phi_2(x)$, in accordance with the Method of Successive Approximations, or would you prefer to differentiate the differential equation two more times?

(g) Explain why the solution $\phi(x)$ is an even function of x whose interval of definition is $(-\infty,+\infty)$.

2. Consider the nonlinear integral equation

$$\phi(x) = 1 + \lambda \int_0^1 xt\,(\phi(t) + \phi^2(t))\,\mathrm{d}t.$$

Every solution to this equation has the form $\phi(x) = 1 + \lambda cx$, where

$$c = \int_0^1 t\,(\phi(t) + \phi^2(t))\,\mathrm{d}t.$$

Thus, the solution set consists of a family of straight lines passing through the point $(0,1)$. By following the procedures shown in Illustrative Example 2, determine all real-valued solutions to this equation. Explain why this equation has no real-valued solutions if $\lambda = 2$ and why it never has any more than two solutions for any value of λ.

3. Consider the nonlinear integral equation

$$\phi(x) = 1 + \lambda \int_0^1 xt\,\left(\phi^2(t) + \phi^3(t)\right)\,\mathrm{d}t.$$

Every solution to this equation has the form $\phi(x) = 1 + \lambda cx$, where

$$c = \int_0^1 t\,\left(\phi^2(t) + \phi^3(t)\right)\,\mathrm{d}t.$$

Thus, the solution set consists of a family of straight lines passing through the point $(0,1)$. Explain why this equation always has at least one real-valued solution for every real value of λ, and why it never has any more than three solutions for any value of λ. Determine all three real-valued solutions if $\lambda = \frac{1}{4}$.

4. Consider the nonlinear integral equation

$$\phi(x) = 1 + \lambda \int_0^1 xt\,\phi(t)\,\mathrm{d}t + \mu \int_0^1 x^2t^2\,\phi^2(t)\,\mathrm{d}t,$$

where $\lambda \neq 0$ and $\mu \neq 0$ are arbitrary parameters. Every solution of this integral equation has the form

$$\phi(x) = 1 + \lambda ax + \mu bx^2,$$

where

$$a = \int_0^1 t\,\phi(t)\,\mathrm{d}t \quad \text{and} \quad b = \int_0^1 t^2\,\phi^2(t)\,\mathrm{d}t.$$

Thus, the solution set consists of a family of parabolas passing through the point $(0,1)$.

(a) By substituting the representation for $\phi(x)$ into the expressions for a and b, show that the two conditions

$$a = \frac{1}{3}\lambda a + \frac{1}{4}\mu b + \frac{1}{2}$$

and

$$b = \frac{1}{5}\lambda^2 a^2 + \frac{1}{3}\lambda a \mu b + \frac{1}{7}\mu^2 b^2 + \frac{1}{2}\lambda a + \frac{2}{5}\mu b + \frac{1}{3}$$

must necessarily be satisfied. (The first equation represents a straight line, and the second equation represents a hyperbola in the ab-plane. The number of solutions varies with the values of λ and μ, since these two curves may be disjoint, or they may intersect in one or more points.)

(b) Let $\lambda = 3$. Use the linear coefficient condition to show that it is necessary for $\mu b = -2$ for a solution to exist. In this case, we have $\phi(x) = 1 + 3ax - 2x^2$. Then use the hyperbolic condition to show that it is necessary for

$$a = \frac{35\mu \pm \sqrt{7}\sqrt{-\mu(10080 + 353\mu)}}{252\,\mu}$$

for a solution to exist. Explain why two real-valued solutions will exist if $-\frac{10080}{353} < \mu < 0$.

Show also that if $\mu = -\frac{10080}{353}$, then $a = \frac{5}{36}$ and $b = \frac{353}{5040}$. Verify that $\phi(x) = 1 + \frac{5}{12}x - 2x^2$ is the unique solution to the integral equation

$$\phi(x) = 1 + 3\int_0^1 xt\,\phi(t)\,\mathrm{d}t - \frac{10080}{353}\int_0^1 x^2 t^2\,\phi^2(t)\,\mathrm{d}t.$$

(c) Suppose that $\lambda \neq 3$. Use the linear condition to show that

$$\mu b = -\frac{2}{3}\left(3 + 2a(\lambda - 3)\right).$$

Multiply the hyperbolic condition by μ and then substitute the above expression for μb into it. The resulting expression can be solved for a in terms of λ and μ alone.

(d) Let $\lambda = 8$ and $\mu = 2$. Determine a and μb. Use these computed values to show that the functions

$$\phi(x) = 1 - \frac{1}{6}\left(55 \pm \sqrt{1169}\right)x + \frac{1}{36}\left(203 \pm 5\sqrt{1169}\right)x^2$$

solve the integral equation

$$\phi(x) = 1 + 8\int_0^1 xt\,\phi(t)\,\mathrm{d}t + 2\int_0^1 x^2t^2\,\phi^2(t)\,\mathrm{d}t.$$

5. Let $P(z) = 2z + 3z^2 + z^3$.

   (a) Consider the nonlinear integral equation of Fredholm type

   $$\phi(x) = 1 + \lambda\int_0^1 P(\phi(t))\,\mathrm{d}t.$$

   Every solution to this equation is a constant of the form $\phi(x) = 1 + \lambda c$, where

   $$c = \int_0^1 P(\phi(t))\,\mathrm{d}t.$$

   How many solutions to this integral equation are possible? Find all real-valued solutions if $\lambda = 1$ and if $\lambda = 5$. Discuss the role that the degree of $P(z)$ plays in answering this question.
   Answer: If $\lambda = 1$, then $\phi(x) \approx -2.769292$, and if $\lambda = 5$, then $\phi(x) \approx -0.144022, -0.621467$, and -2.234510.

   (b) Consider the nonlinear integral equation of Volterra type

   $$\phi(x) = 1 + \lambda\int_0^x P(\phi(t))\,\mathrm{d}t.$$

   Differentiate this integral equation to show that the solution $\phi(x)$ satisfies the nonlinear ordinary differential equation

   $$\phi'(x) = \lambda\left(2\phi(x) + 3\phi^2(x) + \phi^3(x)\right)$$

   together with the initial condition $\phi(0) = 1$. Solve this differential equation, and note the dependence of the domain of the solution upon λ.
   Answer: $\phi(x) = -1 + 2/\sqrt{4 - 3\,\mathrm{e}^{\lambda x}}$.

   (c) Use parts (a) and (b) to investigate what happens when $P(z)$ is an arbitrary polynomial.

6. Use the Laplace transform to solve the nonlinear integral equation

   $$x\phi(x) = \int_0^x \phi(x-t)\,\phi(t)\,\mathrm{d}t.$$

   Is the solution unique in some sense?

Now consider the more complicated equation

$$\phi(x) = f(x) + \lambda \int_0^x \phi(x-t)\,\phi(t)\,\mathrm{d}t.$$

For a solution to exist, it is necessary that $\phi(0) = f(0)$. If so, will this equation always have at least one solution? At least two? What is the maximum number of solutions to this equation?

7. Consider the nonlinear integral equation

$$\phi(x) = f(x) + \lambda \int_a^b G(x,t,\phi(t))\,\mathrm{d}t$$

What necessary conditions would you impose on $f(x)$ and $G(x,t,z)$ to ensure that the solution $\phi(x)$ is bounded? To ensure that $\phi(x) \to 0$ as $x \to +\infty$? To ensure that $\phi(x)$ is infinitely differentiable?

8. Solve each of the following nonlinear integral equations of Fredholm type:

(a) $\phi(x) = \int_0^\pi \sin(x+t)\,\phi^2(t)\,\mathrm{d}t.$

(b) $\phi(x) = \int_0^\pi \cos(x-t)\,\phi^3(t)\,\mathrm{d}t.$

(c) $\phi(x) = \int_0^\pi \cos(x-t)\,\phi^4(t)\,\mathrm{d}t.$

9. Do the following nonlinear integral equations of Volterra type have nontrivial solutions?

(a) $\phi(x) = \int_0^x \sin(x+t)\,\phi^2(t)\,\mathrm{d}t.$

(b) $\phi(x) = \int_0^x \cos(x-t)\,\phi^3(t)\,\mathrm{d}t.$

(c) $\phi(x) = \int_0^x \cos(x-t)\,\phi^4(t)\,\mathrm{d}t.$

10. Adapt the quadrature methods that were described previously in Sect. 2.6.3 for linear integral equations of Fredholm type to nonlinear integral equations.
11. Adapt the quadrature methods that were given previously in Illustrative Example 5 in Sect. 4.4 for linear integral equations of Volterra type to nonlinear integral equations.

Chapter 7
Singular Integral Equations

The theory introduced in previous chapters, especially the Fredholm Theory, was presented under the restrictive assumptions that the kernel was continuous on its domain of definition and that the interval of integration was finite. There is no guarantee that those results or similar ones will hold if the kernel has an infinite discontinuity or if the interval of integration is infinite.

A *singular integral equation* is an equation in which the integral appearing therein is either a convergent improper integral or a divergent improper integral that exists in the sense of the Cauchy principal value. Examples of such integrals include

$$\int_0^x \frac{1}{\sqrt{x-t}}\,\phi(t)\,\mathrm{d}t, \qquad \int_{-1}^{+1} \frac{1}{x-t}\,\phi(t)\,\mathrm{d}t, \quad \text{and} \quad \int_0^{+\infty} \cos(xt)\,\phi(t)\,\mathrm{d}t.$$

A *singular kernel* is a kernel that has an infinite discontinuity in the interior of the interval of integration or at a boundary point of it. According to this definition, the kernels that appear in the first two examples are singular, whereas the kernel that appears in the third example is not.

In this chapter, we will discuss several types of singular integral equations that occur frequently in practice and the techniques that can be used to solve them. Some types of equations, e.g., nonlinear singular integral equations, will not be discussed here. New types of equations are being discovered every day.

In Sect. 7.1, we present several tools of the trade that are indispensible for the comprehension of the material in this chapter.

In Sect. 7.2, we consider singular integral equations with continuous kernels for which the interval of integration is infinite. Many equations of this type can be solved with integral transforms.

In Sect. 7.3, we consider singular integral equations with singular kernels. Since the interval of integration can be either finite or infinite, we consider both of these possibilities. In particular, singular integral equations with the Cauchy kernel and with the Hilbert kernel are considered.

S.M. Zemyan, *The Classical Theory of Integral Equations: A Concise Treatment*,
DOI 10.1007/978-0-8176-8349-8_7, © Springer Science+Business Media, LLC 2012

7.1 Tools of the Trade

In this chapter, the reader should be familiar with the following topics:

- *The evaluation of improper integrals*: There are two types of improper integrals: (1) the integrals for which the integrand possesses an infinite discontinuity within the interval of integration and (2) the integrals for which the interval of integration is infinite. The existence of both types of integrals depends upon the existence of one or more limits. An improper integral is called *convergent* if all of the limits involved exist, and *divergent* if even one of them does not exist.

  (1) If the integrand $f(t)$ is continuous on the interval $[a,c)$ and it has an infinite discontinuity at the right endpoint of the interval, then we define the integral of $f(t)$ on the interval $[a,c)$ to be

$$\int_a^c f(t)\,\mathrm{d}t = \lim_{\varepsilon \downarrow 0} \int_a^{c-\varepsilon} f(t)\,\mathrm{d}t,$$

  provided that the limit exists.

  Similarly, if the integrand has an infinite discontinuity at the left endpoint of the interval, then we define the integral of $f(t)$ on the interval $(c,b]$ to be

$$\int_c^b f(t)\,\mathrm{d}t = \lim_{\varepsilon \downarrow 0} \int_{c+\varepsilon}^b f(t)\,\mathrm{d}t,$$

  provided that the limit exists.

  If the integrand has an infinite discontinuity at an interior point c of the interval $[a,b]$, then we define

$$\int_a^b f(t)\,\mathrm{d}t = \int_a^c f(t)\,\mathrm{d}t + \int_c^b f(t)\,\mathrm{d}t,$$

  provided that both one-sided limits exist independently.

  (2) If the integrand $f(t)$ is continuous on the interval $[c,+\infty)$, then we define the integral of $f(t)$ on this interval to be

$$\int_c^{+\infty} f(t)\,\mathrm{d}t = \lim_{b \to +\infty} \int_c^b f(t)\,\mathrm{d}t,$$

  provided that the limit exists.

  If the integrand is continuous on the interval $(-\infty,c]$, then we define the integral of $f(t)$ on this interval to be

$$\int_{-\infty}^{c} f(t)\,\mathrm{d}t = \lim_{a\to-\infty} \int_{a}^{c} f(t)\,\mathrm{d}t,$$

provided that the limit exists.

If the integrand is continuous on the interval $(-\infty,+\infty)$, then we define the integral of $f(t)$ on this interval to be

$$\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t = \int_{-\infty}^{c} f(t)\,\mathrm{d}t + \int_{c}^{+\infty} f(t)\,\mathrm{d}t,$$

provided that both one-sided limits involved exist independently.

A given integral may display both types of improprieties.

- *The Cauchy principal value of an integral*: In the previous item, we observed that if the integrand has an infinite discontinuity in the interior of the interval of integration or if the interval of integration extended to infinity in both directions, then the existence of the integral depends upon the existence of two limits independently. Even if neither of these two limits exists, a single symmetric limit might still exist.

  For an integral whose integrand $f(t)$ has an infinite discontinuity c in the interior of the interval of integration, we define the *Cauchy principal value of the integral of* $f(t)$ *on the interval* $[a,b]$ to be

$$PV \int_{a}^{b} f(t)\,\mathrm{d}t = \lim_{\varepsilon\downarrow 0} \left(\int_{a}^{c-\varepsilon} f(t)\,\mathrm{d}t + \int_{c+\varepsilon}^{b} f(t)\,\mathrm{d}t \right),$$

  provided that this limit exists.

  For an integral whose interval of integration extends to infinity in both directions, we define the *Cauchy principal value of the integral of* $f(t)$ *on the interval* $(-\infty,+\infty)$ to be

$$PV \int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t = \lim_{a\to\infty} \left(\int_{-a}^{+a} f(t)\,\mathrm{d}t \right),$$

  provided that this limit exists.

  We shall always write PV in front of integrals defined as above. It is important to note that both of these limits are defined as symmetric limits. A non-symmetric limit might also exist, but its value could very well be different.

  Consider the following illustrative example. On the one hand, we have

$$PV \int_{0}^{2} \frac{1}{t-1}\,\mathrm{d}t = \lim_{\varepsilon\varepsilon\downarrow 0} \left(\int_{0}^{1-\varepsilon} \frac{1}{t-1}\,\mathrm{d}t + \int_{1+\varepsilon}^{2} \frac{1}{t-1}\,\mathrm{d}t \right) = 0.$$

On the other hand, we have

$$\int_0^{1-2\varepsilon} \frac{1}{t-1}\,\mathrm{d}t + \int_{1+\varepsilon}^{2} \frac{1}{t-1}\,\mathrm{d}t = \int_{1+\varepsilon}^{1+2\varepsilon} \frac{1}{t-1}\,\mathrm{d}t = \ln 2.$$

There is nothing special about the choice of 2ε here. If we had chosen $k\varepsilon$, then the value of the resulting integral would have been $\ln k$. The point here is that the principal value of an integral is a specific value that is chosen from an infinite number of possibilities.

Another example that illustrates the importance of symmetric limits is the following. On the one hand, we have

$$PV \int_{-\infty}^{+\infty} \frac{2t}{1+t^2}\,\mathrm{d}t = \lim_{a\to\infty} \left(\int_{-a}^{+a} \frac{2t}{1+t^2}\,\mathrm{d}t \right) = 0,$$

since the integrand is odd. On the other hand, we have

$$\int_{-a}^{+2a} \frac{2t}{1+t^2}\,\mathrm{d}t = \int_{+a}^{+2a} \frac{2t}{1+t^2}\,\mathrm{d}t = \ln\left(\frac{1+4a^2}{1+a^2}\right) \to \ln 4,$$

as $a \to \infty$. There is nothing special about the choice of $2a$ here. If we had chosen ka, then the value of the resulting integral would have been $\ln(k^2)$.

- *The beta integral*: There are several equivalent definitions of the beta function $B(p,q)$, all of which are useful. Perhaps the simplest one is

$$B(p,q) = \int_0^1 u^{p-1}\,(1-u)^{q-1}\,\mathrm{d}u,$$

where it is required that p and q are positive for the integral to exist. An elementary change of variables shows that $B(p,q) = B(q,p)$.

If we set $u = \sin^2(\theta)$, then the defining integral becomes

$$B(p,q) = 2\int_0^{\pi/2} \sin^{2p-1}(\theta)\,\cos^{2q-1}(\theta)\,\mathrm{d}\theta.$$

If we set $u = x/(1+x)$, then the defining integral becomes

$$B(p,q) = \int_0^{\infty} \frac{x^{p-1}}{(1+x)^{p+q}}\,\mathrm{d}x.$$

It can be shown that

$$B(p,q) = \frac{\Gamma(p)\,\Gamma(q)}{\Gamma(p+q)}$$

for all choices of $p > 0$ and $q > 0$. In particular, if $p + q = 1$, then we have the relationship

$$B(p, 1-p) = \Gamma(p)\,\Gamma(1-p) = \frac{\pi}{\sin(\pi p)}.$$

The value $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ is deduced by letting $p = \frac{1}{2}$.

- *The Mellin transform*: Suppose that $f(t)$ is continuous on the interval $(0,\infty)$ and that it satisfies the absolute integrability condition

$$\int_0^\infty t^{s-1}\,|f(t)|\,\mathrm{d}t < +\infty$$

for some complex value of $s = \sigma + \mathrm{i}\tau$.

The *Mellin transform of* $f(t)$ is defined by

$$F(s) = \mathcal{M}\{f(t)\} = \int_0^\infty t^{s-1}\,f(t)\,\mathrm{d}t = \lim_{A\to\infty}\int_0^A t^{s-1}\,f(t)\,\mathrm{d}t.$$

If both of the integrability conditions

$$\int_0^1 t^{\sigma_1-1}\,|f(t)|\,\mathrm{d}t < +\infty \quad \text{and} \quad \int_1^\infty t^{\sigma_2-1}\,|f(t)|\,\mathrm{d}t < +\infty$$

hold, then $F(s)$ is analytic in the infinite strip $\Sigma = \{s\colon\ \sigma_1 < \sigma < \sigma_2\}$.

The *inverse Mellin transform of* $F(s)$ is given by

$$f(t) = \mathcal{M}^{-1}\{F(s)\} = \frac{1}{2\pi\mathrm{i}}\int_{c-\mathrm{i}\infty}^{c+\mathrm{i}\infty} t^{-s}\,F(s)\,\mathrm{d}s, \tag{7.1}$$

where $\sigma_1 < c < \sigma_2$. At a value of t where $f(t)$ is discontinuous, the integral converges to the average of the left- and right-hand limits, i.e., $(f(t+) + f(t-))/2$.

The evaluations of both of these Mellin-type integrals often involve the use of the Residue Theorem. Alternately, tables of Mellin transforms[1] are available to assist the student.

[1] See Oberhettinger, F., *Tables of Mellin Transforms*, Springer, New York, 1974.

The *Mellin convolution of* $f(t)$ and $g(t)$ is defined by

$$(f * g)(t) = \int_0^\infty \frac{1}{u} f\left(\frac{t}{u}\right) g(u)\,\mathrm{d}u.$$

If $F(s) = \mathcal{M}\{f(t)\}$ and $G(s) = \mathcal{M}\{g(t)\}$, then $\mathcal{M}\{(f * g)(t)\} = F(s)G(s)$. We also have the similar formula

$$\mathcal{M}\left\{\int_0^\infty f(tu)\,g(u)\,\mathrm{d}u\right\} = F(s)G(1-s), \tag{7.2}$$

which can be derived directly from the definition. The Mellin transform is extremely useful for solving singular integral equations involving integrals of this type.

- *Fourier transforms*: The *Fourier Integral Theorem* is the starting point in any discussion of Fourier Transform pairs. If $f(x)$ and $f'(x)$ are piecewise continuous functions and if $f(x)$ is absolutely integrable on the interval $(-\infty, +\infty)$, then $f(x)$ has the representation

$$f(x) = \frac{1}{\pi} \int_0^{+\infty} \int_{-\infty}^{+\infty} f(t)\,\cos[s(t-x)]\,\mathrm{d}t\,\mathrm{d}s$$

for all values of x for which $f(x)$ is continuous. Furthermore, if x is a value for which $f(x)$ has a jump discontinuity at x, then the integral converges to the average value of the left- and right-hand limits of $f(x)$ at x, i.e., to $(f(x-) + f(x+))/2$.

This double integral has an alternate form that can be written in terms of complex exponentials. Since $\cos\theta = (\mathrm{e}^{+\mathrm{i}\theta} + \mathrm{e}^{-\mathrm{i}\theta})/2$, we also have

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \mathrm{e}^{-\mathrm{i}sx} \int_{-\infty}^{+\infty} \mathrm{e}^{+\mathrm{i}st} f(t)\,\mathrm{d}t\,\mathrm{d}s.$$

We assume that s and t in these integral representations are real variables.

There are many ways to decompose the complex exponential form of Fourier's integral into a pair of Fourier Transforms. A general way of doing so is to write

$$F(s) = \sqrt{\frac{|b|}{(2\pi)^{1-a}}} \int_{-\infty}^{+\infty} f(t)\,\mathrm{e}^{\mathrm{i}bst}\,\mathrm{d}t$$

and

$$f(x) = \sqrt{\frac{|b|}{(2\pi)^{1+a}}} \int_{-\infty}^{+\infty} F(s)\,\mathrm{e}^{-\mathrm{i}bxs}\,\mathrm{d}s,$$

where a and b are real parameters. In this textbook, we choose $a = 0$ and $b = 1$, which is the default choice in the *Mathematica 8.0* software package. It is also the choice in textbooks[2] intended for a more advanced audience. With this choice, the Fourier Transform pair becomes

$$F(s) = \mathcal{F}\{f(t)\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{e}^{\mathrm{i}st} f(t)\,\mathrm{d}t \tag{7.3}$$

and

$$f(x) = \mathcal{F}^{-1}\{F(s)\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-\infty} \mathrm{e}^{-\mathrm{i}xs} F(s)\,\mathrm{d}s. \tag{7.4}$$

Other choices for the Fourier parameters, such as $(-1,+1)$, $(+1,+1)$, $(-1,+1)$, or $(0,-2\pi)$, are commonly used in probability theory, signal processing, or classical physics. Each choice has its advantages.

If $f(x)$ is absolutely integrable, then $F(s)$ is bounded, since

$$|F(s)| \le \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} |f(x)|\,\mathrm{d}x < +\infty.$$

Similarly, if $F(s)$ is absolutely integrable, then $f(x)$ is bounded, since

$$|f(x)| \le \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} |F(s)|\,\mathrm{d}s < +\infty.$$

The Fourier Transform $F(s)$ is continuous even if $f(x)$ is only piecewise continuous. Furthermore, if $F(s)$ exists, then $F(s) \to 0$ as $s \to \pm\infty$.

If $f(t)$ and $g(t)$ are Riemann integrable on every finite interval $[a,b]$ and $|f(t-u)\,g(u)|$ is absolutely integrable for each $t \in (-\infty,+\infty)$, then the *Fourier convolution of $f(t)$ and $g(t)$* is defined to be

$$(f * g)(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(t-u)\,g(u)\,\mathrm{d}u,$$

or equivalently, with an elementary change of variable,

$$(f * g)(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(u)\,g(t-u)\,\mathrm{d}u.$$

[2] See Porter, D. and Stirling, D., *Integral Equations: A Practical Treatment, from Spectral Theory to Applications*, Cambridge University Press, Cambridge, 1990. For an alternate choice, see Andrews, L.C. and Shivamoggi, B.K., *Integral Transforms for Engineers and Applied Mathematicians*, MacMillan Publishing Company, New York, 1988.

The integrability condition is fulfilled, for example, if $f(t)$ and $g(t)$ are square integrable, since an application of the Cauchy–Schwarz inequality gives

$$\int_{-\infty}^{+\infty} |f(t-u)\,g(u)|\,\mathrm{d}u \le \left(\int_{-\infty}^{+\infty} |f(u)|^2\,\mathrm{d}u\right)^{1/2} \left(\int_{-\infty}^{+\infty} |g(u)|^2\,\mathrm{d}u\right)^{1/2}.$$

The Fourier Transform is extremely useful for solving singular integral equations involving integrals of convolution type. In order to do so, we must be able to compute the Fourier Transform of a convolution integral. Let $F(s) = \mathcal{F}\{f(t)\}$ and $G(s) = \mathcal{F}\{g(t)\}$. Then the Fourier Convolution Theorem states that $\mathcal{F}\{(f * g)(t)\} = F(s)G(s)$. It is often stated alternately in the form

$$\int_{-\infty}^{+\infty} \mathrm{e}^{-\mathrm{i}st} F(s)\,G(s)\,\mathrm{d}s = \int_{-\infty}^{+\infty} f(u)\,g(t-u)\,\mathrm{d}u. \tag{7.5}$$

In particular, if $t = 0$ and $g(-u) = \overline{f(u)}$, then $G(s) = \overline{F(s)}$ and

$$\int_{-\infty}^{+\infty} |F(s)|^2\,\mathrm{d}s = \int_{-\infty}^{+\infty} |f(t)|^2\,\mathrm{d}t. \tag{7.6}$$

This relationship between $f(t)$ and its Fourier Transform $F(s)$ is called Parseval's relation. It follows that $f(t)$ is square integrable if and only if $F(s)$ is square integrable.

Quite often, it is necessary to compute the Fourier Transforms of the derivatives of $f(t)$ in terms of the Fourier Transform of $f(t)$. The formula

$$\mathcal{F}\{f^{(n)}(t)\} = (-\mathrm{i}s)^n F(s) \tag{7.7}$$

holds, under the necessary assumption that $f^{(n)}(t)$ is piecewise smooth and absolutely integrable, in addition to the reasonable assumptions that $f(t), f'(t), \ldots, f^{(n-1)}(t)$ are continuous everywhere, vanish as $t \to \pm\infty$, and are absolutely integrable on the interval $(-\infty, +\infty)$.

If $f(t)$ is defined and integrable only on the interval $[0, +\infty)$, then the Fourier Transform of a symmetrical extension of it can still be computed.

The *even extension of* $f(t)$ to $(-\infty, +\infty)$ is defined by setting $f_E(t) = f(|t|)$, and the *odd extension of* $f(t)$ to $(-\infty, +\infty)$ is defined by setting $f_O(t) = \mathit{signum}(t)\,f(|t|)$, where $\mathit{signum}(t)$ is equal to $+1$ if t is positive and -1 if t is negative.

If $f(t)$ is even, then the Fourier integral representation assumes the simplified form

$$f(x) = \frac{2}{\pi} \int_0^{\infty} \int_0^{\infty} \cos(xs)\cos(st)\,f(t)\,\mathrm{d}t\,\mathrm{d}s.$$

The Fourier Transform (7.3) becomes the *Fourier Cosine Transform*

$$F_{\mathrm{C}}(s) = \mathcal{F}_{\mathrm{C}}\{f(t)\} = \sqrt{\frac{2}{\pi}} \int_0^\infty \cos(st)\, f(t)\, \mathrm{d}t \quad (s > 0) \tag{7.8}$$

and $\mathcal{F}\{f(t)\} = \mathcal{F}_{\mathrm{C}}\{f(t)\}$.

The Inverse Fourier Transform (7.4) becomes the *Inverse Fourier Cosine Transform*

$$f(x) = \mathcal{F}_{\mathrm{C}}^{-1}\{F_{\mathrm{C}}(s)\} = \sqrt{\frac{2}{\pi}} \int_0^\infty \cos(xs)\, F_{\mathrm{C}}(s)\, \mathrm{d}s. \quad (x > 0) \tag{7.9}$$

If $f(t)$ is odd, then the Fourier integral representation assumes the simplified form

$$f(x) = \frac{2}{\pi} \int_0^\infty \int_0^\infty \sin(xs)\, \sin(st)\, f(t)\, \mathrm{d}t\, \mathrm{d}s.$$

The Fourier Transform (7.3) leads to the *Fourier Sine Transform*

$$F_{\mathrm{S}}(s) = \mathcal{F}_{\mathrm{S}}\{f(t)\} = \sqrt{\frac{2}{\pi}} \int_0^\infty \sin(st)\, f(t)\, \mathrm{d}t \quad (s > 0) \tag{7.10}$$

and $\mathcal{F}\{f(t)\} = \mathrm{i}\mathcal{F}_{\mathrm{S}}\{f(t)\}$.

The Inverse Fourier Transform (7.4) leads to the *Inverse Fourier Sine Transform*

$$f(x) = \mathcal{F}_{\mathrm{S}}^{-1}\{F_{\mathrm{S}}(s)\} = \sqrt{\frac{2}{\pi}} \int_0^\infty \sin(xs)\, F_{\mathrm{C}}(s)\, \mathrm{d}s \quad (x > 0) \tag{7.11}$$

7.2 Equations with Continuous Kernels

It is not our intent in this section to give a complete introduction to the theory, properties, and uses of the Mellin, Fourier, and Laplace integral transforms. However, since it is our intent to illustrate how these transforms can be used to solve singular integral equations of certain types, the introduction of some of their properties along the way will be essential.

In this section, we will consider singular integral equations with continuous kernels for which the interval of integration is infinite. The intervals of interest are $(-\infty, a)$, $(a, +\infty)$, $(-\infty, x)$, $(x, +\infty)$, and $(-\infty, +\infty)$. The limit a is a constant, but it is usually taken to be 0 or 1.

A wide variety of integral equations fall into this category. The kernels that appear most frequently in practice are of the form $K(x,t) = k(x-t)$, $J(x,t) = j(xt)$, $M(x,t) = m(x/t)$, $E(x,t) = \mathrm{e}^{\mathrm{i}xt}$, $C(x,t) = \cos(xt)$, or $S(x,t) = \sin(xt)$.

Many of the techniques that are used to solve singular integral equations require the use of integral transforms. The choice of transform (Mellin, Fourier, or Laplace) to be used depends upon the form of the kernel and the interval of integration.

The techniques that will be used here are similar to the Laplace transform method that was previously used to solve a Volterra integral equation of the second kind of the form

$$\phi(x) = f(x) + \lambda \int_0^x k(x-t)\,\phi(t)\,\mathrm{d}t.$$

We recall from Sect. 4.3 that the Laplace transform converted the integral equation into an algebraic equation that could be solved for the transform $\Phi(s)$ of the solution $\phi(x)$ in terms of the transform $F(s)$ of the inhomogeneous term $f(x)$ and the transform $K(s)$ of the function $k(x)$, provided that all of the transforms exist. The solution $\phi(x)$ could then be recovered from $\Phi(s)$.

The singular integral equations

$$\phi(x) = \lambda \int_0^{+\infty} K(x,t)\,\phi(t)\,\mathrm{d}t \quad \text{and} \quad \phi(x) = \lambda \int_{-\infty}^{+\infty} K(x,t)\,\phi(t)\,\mathrm{d}t$$

are of interest. The values λ for which nontrivial solutions to these equations exist are called *eigenvalues of the kernel*. The solutions are called *eigenfunctions*, in keeping with the terminology introduced in earlier chapters. We will see that the Fourth Fredholm Theorem need not hold for equations of this type.

We shall also be concerned with singular integral equations of the first kind of the forms

$$f(x) = \lambda \int_0^{+\infty} K(x,t)\,\phi(t)\,\mathrm{d}t \quad \text{and} \quad f(x) = \lambda \int_{-\infty}^{+\infty} K(x,t)\,\phi(t)\,\mathrm{d}t,$$

and singular integral equations of the second kind of the forms

$$\phi(x) = f(x) + \lambda \int_0^{+\infty} K(x,t)\,\phi(t)\,\mathrm{d}t$$

and

$$\phi(x) = f(x) + \lambda \int_{-\infty}^{+\infty} K(x,t)\,\phi(t)\,\mathrm{d}t.$$

7.2.1 Equations Solved with the Mellin Transform

Consider the singular integral equation of the second kind

$$\phi(x) = f(x) + \lambda \int_0^\infty \frac{1}{t} k\left(\frac{x}{t}\right) \phi(t)\,\mathrm{d}t.$$

Since the integral is given in terms of the Mellin convolution of $k(x)$ and $\phi(x)$, an application of the Mellin Transform to both sides of the equation yields

$$\Phi(s) = F(s) + \lambda K(s)\Phi(s),$$

provided that these transforms exist for common values of s in some infinite strip $\Sigma = \{s\colon \sigma_1 < s < \sigma_2\}$. By solving for $\Phi(s)$, we obtain

$$\Phi(s) = \frac{F(s)}{1 - \lambda K(s)},$$

provided that $\lambda K(s) \neq 1$. If $\Phi(s)$ is recognized in a table of Mellin transforms, then the problem is essentially solved. If not, then it is necessary to apply the Inverse Mellin Transform which yields the formal solution

$$\phi(x) = \frac{1}{2\pi \mathrm{i}} \int_{c-\mathrm{i}\infty}^{c+\mathrm{i}\infty} \frac{F(s)}{1 - \lambda K(s)} x^{-s}\,\mathrm{d}s,$$

where $\sigma_1 < c < \sigma_2$. This integral is usually evaluated with the assistance of the Residue Theorem after an appropriate contour is chosen.

The singular integral equation of the first kind

$$f(x) = \int_0^\infty k(xt)\,\phi(t)\,\mathrm{d}t$$

can also be formally solved by applying the Mellin Transform to both sides of the equation. After doing so, we obtain $F(s) = K(s)\,\Phi(1-s)$, or equivalently $F(1-s) = K(1-s)\Phi(s)$, after replacing s with $1-s$. We must, of course, assume that the transforms $F(1-s)$ and $K(1-s)$ exist for values of s in some common infinite strip $\Sigma = \{s\colon \sigma_1 < s < \sigma_2\}$. If $\Phi(s)$ is recognized in a table of Mellin Transforms, then the problem is essentially solved. An application of the Inverse Mellin Transform yields the formal solution

$$\phi(x) = \frac{1}{2\pi \mathrm{i}} \int_{c-\mathrm{i}\infty}^{c+\mathrm{i}\infty} \frac{F(1-s)}{K(1-s)} x^{-s}\,\mathrm{d}s,$$

for a suitable value of c.

Fox's singular integral equation

$$\phi(x) = f(x) + \int_0^\infty k(xt)\,\phi(t)\,\mathrm{d}t$$

can also be formally solved by applying the Mellin Transform. After doing so, we obtain

$$\Phi(s) = F(s) + K(s)\,\Phi(1-s).$$

After replacing s by $1-s$, we obtain

$$\Phi(1-s) = F(1-s) + K(1-s)\,\Phi(s).$$

Upon eliminating $\Phi(1-s)$ from these equations, we can easily solve for $\Phi(s)$ to obtain

$$\Phi(s) = \frac{F(s) + K(s)\,F(1-s)}{1 - K(s)\,K(1-s)},$$

which then results in the formal solution

$$\phi(x) = \frac{1}{2\pi\mathrm{i}} \int_{c-\mathrm{i}\infty}^{c+\mathrm{i}\infty} \frac{F(s) + K(s)\,F(1-s)}{1 - K(s)\,K(1-s)}\,x^{-s}\,\mathrm{d}s,$$

for a suitable value of c.

Illustrative Examples

- *Example 1*: Show that the singular integral equation

$$\phi(s) = \lambda \int_0^\infty \mathrm{e}^{-st}\,\phi(t)\,\mathrm{d}t$$

has nontrivial solutions.

After scanning a standard table of Laplace transforms, we find that

$$\mathcal{L}\{t^{-1/2}\} = \sqrt{\pi}\,s^{-1/2},$$

or, equivalently, that

$$s^{-1/2} = \frac{1}{\sqrt{\pi}} \int_0^\infty \mathrm{e}^{-st}\,t^{-1/2}\,\mathrm{d}t. \tag{7.12}$$

Thus, with $\lambda = 1/\sqrt{\pi}$ and $\phi(t) = t^{-1/2}$, there is at least one nontrivial solution to this equation.

More generally, if $0 < p < 1$, then

$$\mathcal{L}\{t^{-p}\} = \Gamma(1-p)\,s^{p-1} \quad \text{and} \quad \mathcal{L}\{t^{p-1}\} = \Gamma(p)\,s^{-p}.$$

If we set

$$\phi_p^+(t) = \sqrt{\Gamma(p)}\,t^{-p} + \sqrt{\Gamma(1-p)}\,t^{p-1} \tag{7.13}$$

and

$$\phi_p^-(t) = \sqrt{\Gamma(p)}\,t^{-p} - \sqrt{\Gamma(1-p)}\,t^{p-1}, \tag{7.14}$$

then we have

$$\mathcal{L}\{\phi_p^+(t)\} = \sqrt{\Gamma(p)\,\Gamma(1-p)}\,\phi_p^+(s) = +\sqrt{\frac{\pi}{\sin(p\pi)}}\,\phi_p^+(s)$$

and

$$\mathcal{L}\{\phi_p^-(t)\} = -\sqrt{\Gamma(p)\,\Gamma(1-p)}\,\phi_p^-(s) = -\sqrt{\frac{\pi}{\sin(p\pi)}}\,\phi_p^-(s).$$

Equivalently, we have

$$\phi_p^{\pm}(s) = \pm\lambda_p \int_0^{\infty} \mathrm{e}^{-st}\,\phi_p^{\pm}(t)\,\mathrm{d}t,$$

with $\pm\lambda_p = \pm\sqrt{\sin(p\pi)/\pi}$.

We have discovered in this example that *every value of* λ in the intervals $(-1/\sqrt{\pi}, 0)$ and $(0, +1\sqrt{\pi}]$ is an eigenvalue of the Laplace kernel. The endpoint $-1/\sqrt{\pi}$ was omitted, since $\phi_{1/2}^-(t) \equiv 0$.[3]

Note that the eigenfunctions $\phi_p^{\pm}(t)$ are continuous on the interval $(0, +\infty)$, but that they have an infinite discontinuity at $t = 0$. Note also that these eigenfunctions are not square integrable on the interval $(0, +\infty)$.

- *Example 2*: Consider the singular integral equation

$$\frac{1}{1+x} = \int_0^{+\infty} \mathrm{e}^{-xt}\,\phi(t)\,\mathrm{d}t.$$

Note that

$$\mathcal{M}\left\{\frac{1}{1+x}\right\} = \frac{\pi}{\sin(\pi s)} = \Gamma(s)\,\Gamma(1-s) \quad \text{and that} \quad \mathcal{M}\left\{\mathrm{e}^{-x}\right\} = \Gamma(s).$$

[3]This example is discussed in greater detail in Widder, D.V., *The Laplace Transform*, Princeton University Press, Princeton, 1946, pp. 390–391.

An application of the Mellin convolution formula (7.2) gives

$$\Gamma(s)\,\Gamma(1-s) = \Gamma(s)\,\Phi(1-s),$$

or, equivalently, $\Phi(s) = \Gamma(s)$. Hence, $\phi(x) = \mathrm{e}^{-x}$, and the solution is easily verified.

7.2.2 Equations Solved with Fourier Transforms

The singular integral equation of the first kind

$$f(x) = \int_{-\infty}^{+\infty} k(x-t)\,\phi(t)\,\mathrm{d}t$$

can be formally solved by applying the Fourier Transform to it. Since the kernel $k(x-t)$ is of convolution type, a straightforward application of the convolution theorem yields

$$F(s) = \sqrt{2\pi}\,K(s)\,\Phi(s),$$

where $F(s) = \mathcal{F}\{f(x)\}$, $K(s) = \mathcal{F}\{k(x)\}$, and $\Phi(s) = \mathcal{F}\{\phi(x)\}$, provided that these transforms exist. If

$$\Phi(s) = \frac{1}{\sqrt{2\pi}}\,\frac{F(s)}{K(s)}$$

is recognized in a table of Fourier Transforms, then the equation is essentially solved. Formally, the solution $\phi(x)$ has the integral representation

$$\phi(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \mathrm{e}^{-\mathrm{i}xs}\,\frac{F(s)}{K(s)}\,\mathrm{d}s,$$

by applying the Inverse Fourier Transform to $\Phi(s)$.

In view of Parseval's relation (7.6), the solution $\phi(x)$ is square integrable if and only if its transform $\Phi(s)$ is square integrable. For this to occur, it is necessary to assume at the outset that $f(x)$ (and hence $F(s)$) is square integrable and $k(x)$ is absolutely integrable. We do not exclude the possibility that the integral equation might have solutions that are not square integrable.

The Fredholm singular integral equation of the second kind

$$\phi(x) = f(x) + \lambda \int_{-\infty}^{+\infty} k(x-t)\,\phi(t)\,\mathrm{d}t$$

can also be formally solved by applying the Fourier Transform to it. A straightforward application of the convolution theorem yields

$$\Phi(s) = F(s) + \lambda \sqrt{2\pi}\, K(s)\, \Phi(s),$$

where $F(s) = \mathcal{F}\{f(x)\}$, $K(s) = \mathcal{F}\{k(x)\}$, and $\Phi(s) = \mathcal{F}\{\phi(x)\}$, provided that these transforms exist. The transform of the solution has the representation

$$\Phi(s) = \frac{F(s)}{1 - \lambda \sqrt{2\pi}\, K(s)},$$

where $1 - \lambda \sqrt{2\pi}\, K(s) \neq 0$ for all real values of s. If $\Phi(s)$ is recognized in a table of Fourier Transforms, then the equation is essentially solved. Formally, the solution $\phi(x)$ has the integral representation

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{e}^{-\mathrm{i}xs}\, \frac{F(s)}{1 - \lambda \sqrt{2\pi}\, K(s)}\, \mathrm{d}s \quad (-\infty < x < +\infty)$$

by applying the Inverse Fourier Transform to $\Phi(s)$. This representation is valid if $f(x)$ is square integrable and $k(x)$ is absolutely integrable. If $1 - \lambda \sqrt{2\pi}\, K(s) = 0$ for some real value of s, then the integral equation might not have a solution that is absolutely integrable.

Another representation for the solution is possible. Recall that

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-\infty} \mathrm{e}^{-\mathrm{i}xs}\, F(s)\, \mathrm{d}s.$$

By subtracting the given representations for $\phi(x)$ and $f(x)$, we obtain

$$\phi(x) = f(x) + \lambda \int_{-\infty}^{+\infty} \mathrm{e}^{-\mathrm{i}xs} \left(\frac{F(s)\, K(s)}{1 - \lambda \sqrt{2\pi}\, K(s)} \right) F(s)\, \mathrm{d}s.$$

If $g(x;\lambda)$ is a function for which

$$\Gamma(s;\lambda) = \mathcal{F}\{g(x;\lambda)\} = \frac{F(s)\, K(s)}{1 - \lambda \sqrt{2\pi}\, K(s)},$$

then the solution can be alternately written in the form

$$\phi(x) = f(x) + \lambda \int_{-\infty}^{+\infty} g(x - t;\lambda)\, f(t)\, \mathrm{d}t,$$

in view of relation (7.5).

Illustrative Examples

- *Example 1*: In this example, we illustrate the various types of integrability in relation to the Fourier Transform. Let

$$f(t) = \begin{cases} 0 & \text{if } -\infty < t < -1 \\ \frac{1}{2} & \text{if } t = -1 \\ 1 & \text{if } -1 < t < +1 \\ \frac{1}{2} & \text{if } t = +1 \\ 0 & \text{if } +1 < t < +\infty. \end{cases}$$

Then $f(t)$ is an even step function that is obviously integrable, absolutely integrable, and square integrable. Its Fourier Transform exists and it is easily computed to be

$$\begin{aligned} F(s) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{e}^{\mathrm{i}st} f(t)\,\mathrm{d}t \\ &= \frac{1}{\sqrt{2\pi}} \int_{-1}^{+1} \mathrm{e}^{\mathrm{i}st} f(t)\,\mathrm{d}t \\ &= \frac{1}{\sqrt{2\pi}} \frac{\mathrm{e}^{it} - \mathrm{e}^{-it}}{is} \\ &= \sqrt{\frac{2}{\pi}} \frac{\sin s}{s}. \end{aligned}$$

Thus, $F(s)$ is also even. The value $F(0) = \sqrt{2/\pi}$ is defined as usual by continuous extension. If we apply the Inverse Fourier Transform to $F(s)$, we have

$$\begin{aligned} f(t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-\infty} \mathrm{e}^{-\mathrm{i}st} F(s)\,\mathrm{d}s \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-\infty} \mathrm{e}^{-\mathrm{i}st} \sqrt{\frac{2}{\pi}} \frac{\sin s}{s}\,\mathrm{d}s \\ &= \frac{2}{\pi} \int_{0}^{+\infty} \cos(st) \frac{\sin s}{s}\,\mathrm{d}s. \end{aligned}$$

In particular, if $t = 0$, then $f(0) = 1$, so that

$$\int_0^{+\infty} \frac{\sin s}{s}\,\mathrm{d}s = \frac{\pi}{2},$$

a well-known result. Thus, $F(s)$ is integrable on $(-\infty, +\infty)$.

Although $F(s)$ is integrable, it is not absolutely integrable. Since

$$\int_{n\pi}^{(n+1)\pi} \left|\frac{\sin s}{s}\right| \mathrm{d}s \geq \frac{1}{(n+1)\pi} \int_{n\pi}^{(n+1)\pi} |\sin s|\,\mathrm{d}s = \frac{2}{(n+1)\pi}$$

for every $n \geq 0$, we have

$$\int_0^{(N+1)\pi} \left|\frac{\sin s}{s}\right| \mathrm{d}s \geq \frac{2}{\pi} \sum_{n=0}^{N} \frac{1}{n+1} \to +\infty$$

as $N \to +\infty$.

However, an application of Parseval's relation (7.6) shows that $F(s)$ is square integrable, since $f(t)$ is square integrable. An interesting integral value can easily be derived from this observation. On the one hand, we have

$$\int_{-\infty}^{+\infty} |f(t)|^2\,\mathrm{d}t = \int_{-1}^{+1} 1^2\,\mathrm{d}t = 2.$$

On the other hand, we have

$$\begin{aligned}\int_{-\infty}^{+\infty} |F(s)|^2\,\mathrm{d}s &= \int_{-\infty}^{+\infty} \left|\sqrt{\frac{2}{\pi}}\,\frac{\sin s}{s}\right|^2 \mathrm{d}s \\ &= \frac{4}{\pi} \int_0^{+\infty} \frac{\sin^2 s}{s^2}\,\mathrm{d}s.\end{aligned}$$

By setting these two integrals equal, we obtain

$$\int_0^{+\infty} \frac{\sin^2 s}{s^2}\,\mathrm{d}s = \frac{\pi}{2}.$$

- *Example 2*: Show that the singular integral equation

$$\phi(s) = \lambda \int_0^{\infty} \cos(st)\,\phi(t)\,\mathrm{d}t$$

has nontrivial solutions.

Let $f(t)$ be any function for which a Fourier Cosine Transform pair exists. After appropriately replacing the dummy variables in the Fourier Cosine Transform pair, we have

$$F_{\mathrm{C}}(x) = \sqrt{\frac{2}{\pi}} \int_0^\infty \cos(xt)\, f(t)\, \mathrm{d}t$$

and

$$f(x) = \sqrt{\frac{2}{\pi}} \int_0^\infty \cos(xt)\, F_{\mathrm{C}}(t)\, \mathrm{d}t.$$

If we add these two equations, then we obtain

$$\phi(x) = \sqrt{\frac{2}{\pi}} \int_0^\infty \cos(xt)\, \phi(t)\, \mathrm{d}t,$$

where $\phi(x) = f(x) + F_{\mathrm{C}}(x)$. Thus, $\phi(x)$ is an eigenfunction corresponding to the eigenvalue $\lambda = \sqrt{2/\pi}$.

For example, if $f(t) = \mathrm{e}^{-a|t|}$, where $a > 0$, then

$$F_{\mathrm{C}}(s) = \mathcal{F}_{\mathrm{C}}\{f(t)\} = \sqrt{\frac{2}{\pi}} \int_0^\infty \cos(st)\, \mathrm{e}^{-at}\, \mathrm{d}t = \sqrt{\frac{2}{\pi}}\, \frac{a}{a^2 + s^2}.$$

Consequently,

$$\phi(x) = \mathrm{e}^{-ax} + \sqrt{\frac{2}{\pi}}\, \frac{a}{a^2 + x^2}$$

is an eigenfunction corresponding to the eigenvalue $\sqrt{2/\pi}$ for every $a > 0$. The functions $f(t)$, $F_{\mathrm{C}}(s)$, and $\phi(x)$ are integrable and square integrable on the interval $(-\infty, +\infty)$.

- *Example 3*: In Sect. 2.4, we considered the Lalesco–Picard integral equation

$$\phi(x) = \lambda \int_{-\infty}^{+\infty} \mathrm{e}^{-|x-t|}\, \phi(t)\, \mathrm{d}t.$$

It was solved by converting it to an ordinary differential equation of the second order with constant coefficients. By doing so, we showed that there were two eigenfunctions associated with each value of $\lambda \in (0, +\infty)$. Recall that there are three possibilities:

  - If $0 < \lambda < \frac{1}{2}$, then the two solutions are $\phi_1(x;\lambda) = \sinh(\mu x)$ and $\phi_2(x;\lambda) = \cosh(\mu x)$, where $\mu^2 = 1 - 2\lambda$.
  - If $\lambda = \frac{1}{2}$, then the two solutions are $\phi_1(x;\frac{1}{2}) = 1$ and $\phi_2(x;\frac{1}{2}) = x$.

– If $\frac{1}{2} < \lambda < +\infty$, then the two solutions are $\phi_1(x;\lambda) = \sin(\mu x)$ and $\phi_2(x;\lambda) = \cos(\mu x)$, where $\mu^2 = 2\lambda - 1$.

With the same kernel, we now consider the singular integral equation of the first kind

$$f(x) = \int_{-\infty}^{+\infty} \mathrm{e}^{-|x-t|}\,\phi(t)\,\mathrm{d}t,$$

and we illustrate how the Fourier Transform can be used to solve it. Since the kernel is a convolution kernel and

$$\mathcal{F}\left\{\mathrm{e}^{-|t|}\right\} = \sqrt{\frac{2}{\pi}}\,\frac{1}{s^2+1},$$

an application of the Fourier convolution theorem directly produces the transformed equation

$$\Phi(s) = \frac{1}{2}F(s) + \frac{1}{2}s^2 F(s).$$

From formula (7.7), we know that $\mathcal{F}\{-f''(t)\} = s^2 F(s)$. Hence, we have the solution

$$\phi(x) = \frac{1}{2}f(x) - \frac{1}{2}f''(t),$$

which is valid under the assumptions that $f(x)$ and $f'(x)$ are continuous and absolutely integrable on the interval $(-\infty,+\infty)$, that they vanish as $t \to \pm\infty$, and that $f''(t)$ is piecewise continuous and absolutely integrable on the interval $(-\infty,+\infty)$.

The inhomogeneous singular integral equation of the second kind

$$\phi(x) = f(x) + \lambda \int_{-\infty}^{+\infty} \mathrm{e}^{-|x-t|}\,\phi(t)\,\mathrm{d}t$$

can also be solved with the Fourier Transform. For the purpose of this example, we will assume that $0 < \lambda < \frac{1}{2}$, so that $0 < 1 - 2\lambda < 1$. After applying the Fourier Transform and simplifying the result, we obtain

$$\Phi(s) = F(s) + \frac{2\lambda}{s^2+1}\,\Phi(s),$$

which can be rearranged as

$$\Phi(s) = F(s) + \lambda\left(\frac{2}{s^2 + \left(\sqrt{1-2\lambda}\right)^2}\right)F(s).$$

Note that the second term on the right-hand side of this equation is the product of two Fourier Transforms. Consequently, after an application of the Inverse Fourier

Transform, the solution to the integral equation can be written in terms of a convolution integral as

$$\phi(x) = f(x) + \frac{\lambda}{\sqrt{1-2\lambda}} \int_{-\infty}^{+\infty} \mathrm{e}^{-\sqrt{1-2\lambda}\,|x-t|} f(t)\,\mathrm{d}t.$$

7.2.3 Equations of Volterra Type

Volterra singular integral equations of the second kind have either of the forms

$$\phi(x) = f(x) + \lambda \int_{x}^{\infty} K(x,t)\,\phi(t)\,\mathrm{d}t$$

or

$$\phi(x) = f(x) + \lambda \int_{-\infty}^{x} K(x,t)\,\phi(t)\,\mathrm{d}t,$$

where it is assumed that the Volterra kernel $K(x,t)$ vanishes in a half-plane.

Suppose that $f(x)$ is continuous and bounded. If the kernel $K(x,t)$ satisfies suitable integrability conditions, then the Method of Successive Approximations can be employed to produce unique solutions to these equations in the respective forms

$$\phi(x) = f(x) + \lambda \int_{x}^{\infty} R(x,t;\lambda)\,f(t)\,\mathrm{d}t$$

or

$$\phi(x) = f(x) + \lambda \int_{-\infty}^{x} R(x,t;\lambda)\,f(t)\,\mathrm{d}t,$$

where $R(x,t;\lambda)$ is the resolvent kernel constructed from the iterations of $K(x,t)$ on each of the respective intervals of integration.

If the kernel $K(x,t)$ is separable or is a difference kernel, then some of the techniques that have been previously introduced in this text may be successfully employed here.

Illustrative Examples

- *Example 1*: If the kernel is separable, then a singular integral equation can often be converted to an ordinary differential equation or to a linear system of ordinary differential equations.

Consider the Volterra singular integral equation

$$\phi(x) = \mathrm{e}^x + 2\int_{-\infty}^{x} \mathrm{e}^{-3(x-t)}\,\phi(t)\,\mathrm{d}t.$$

If we multiply this integral equation by e^{3x} and differentiate the result, we obtain the first-order linear equation

$$\phi'(x) + \phi(x) = 4\mathrm{e}^x$$

after performing the usual simplifications. The solutions to this differential equation are subject only to the initial condition that $\phi'(0) + \phi(0) = 4$. Upon solving this equation, we obtain the solution

$$\phi(x) = (\phi(0) - 2)\,\mathrm{e}^{-x} + 2\mathrm{e}^x.$$

- *Example 2*: Volterra singular integral equations of the second kind

$$\phi(x) = f(x) + \int_{x}^{+\infty} k(x-t)\,\phi(t)\,\mathrm{d}t,$$

whose kernel is a convolution or difference kernel, can be solved with the Laplace transform, although the solution may not be unique.

The essential transform formula is given by

$$\mathcal{L}\left\{\int_{x}^{+\infty} K(x-t)\,\phi(t)\,\mathrm{d}t\right\} = K(-s)\,\Phi(s), \tag{7.15}$$

where

$$K(-s) = \int_{0}^{+\infty} k(-x)\,\mathrm{e}^{sx}\,\mathrm{d}x \quad \text{and} \quad \Phi(s) = \mathcal{L}\{\phi(x)\}.$$

To illustrate this procedure, consider the integral equation

$$\phi(x) = 3\,\mathrm{e}^{-x} + 2\int_{x}^{+\infty} \mathrm{e}^{x-t}\,\phi(t)\,\mathrm{d}t.$$

Since $k(x) = \mathrm{e}^x$, $K(-s) = 1/(1-s)$. After transforming the integral equation, we find after some simplification that

$$\Phi(s) = \frac{3}{s+1} - \frac{6}{(s+1)^2},$$

from which we conclude that

$$\phi(x) = 3\,\mathrm{e}^{-x} - 6x\mathrm{e}^{-x}.$$

However, this solution is not unique. Suppose that there were two different solutions, say $\phi_1(x)$ and $\phi_2(x)$. If we set $\delta(x) = \phi_1(x) - \phi_2(x)$, then $\delta(x)$ must satisfy the equation

$$\delta(x) = 2 \int_x^{+\infty} \mathrm{e}^{x-t}\, \delta(t)\, \mathrm{d}t.$$

After converting this integral equation to a differential equation, we obtain $\delta'(x) + \delta(x) = 0$. Hence, $\delta(x) = c\,\mathrm{e}^{-x}$, where c is an arbitrary constant. It follows that the most general solution to the integral equation has the form

$$\phi(x) = \phi(0)\,\mathrm{e}^{-x} - 6x\mathrm{e}^{-x}.$$

If the integral equation had been converted to a differential equation by following the procedure outlined in the previous example, then this general solution would have been obtained directly.

Sect. 7.2 Exercises

1. Evaluate the integral
$$\int_{-1}^{+1} (1+x)^{p-1}(1-x)^{q-1}\, \mathrm{d}x$$
in terms of the beta function.
2. For $0 < \alpha < +\infty$, let $D(\alpha)$ denote the domain in the first quadrant bounded by the x-axis, the y-axis, and the curve $x^\alpha + y^\alpha = 1$.
   (a) Show that the area $A(\alpha)$ of $D(\alpha)$ is given by
   $$A(\alpha) = \frac{1}{2\alpha} B\left(\frac{1}{\alpha}, \frac{1}{\alpha}\right).$$
   For example, if $\alpha = 1$, then $A(1) = \frac{1}{2}$, and if $\alpha = 2$, then $A(2) = \frac{\pi}{4}$.
   (b) Show that
   $$\iint_{D(\alpha)} \frac{1}{1-xy}\, \mathrm{d}y\, \mathrm{d}x = \frac{1}{2\alpha} \sum_{n=1}^{\infty} \frac{1}{n} B\left(\frac{n}{\alpha}, \frac{n}{\alpha}\right).$$
3. Show that the Laplace kernel $L(x,t) = \mathrm{e}^{-xt}$ is not square integrable on $Q(0, +\infty)$.
4. Does there exist a function for which $\mathcal{M}\{f(x)\} = 1$?
   Hint: Revisit the Inverse Mellin Transform (7.1).

5. Let

$$F_{\mathrm{M}}(x) = \mathcal{M}\{f(t)\} = \int_0^\infty t^{x-1} f(t)\,\mathrm{d}t$$

and

$$F_{\mathrm{L}}(x) = \mathcal{L}\{f(t)\} = \int_0^\infty \mathrm{e}^{-xt} f(t)\,\mathrm{d}t.$$

(a) Show that

$$F_{\mathrm{LM}}(s) = \mathcal{M}\{F_{\mathrm{L}}(x)\} = \Gamma(s) F_{\mathrm{M}}(1-s).$$

You will need to interchange the order of integration and use the definition of the Γ-function.

(b) Try to compute $F_{\mathrm{ML}}(s) = \mathcal{L}\{F_{\mathrm{M}}(x)\}$. Do these integrals exist?

(c) Does $F_{\mathrm{LM}}(s) = F_{\mathrm{ML}}(s)$ under any circumstances? That is, do the operators $\mathcal{L}$ and $\mathcal{M}$ commute?

6. Does the singular integral equation

$$\phi(s) = \lambda \int_0^\infty \sin(st)\,\phi(t)\,\mathrm{d}t$$

have any nontrivial solutions?

7. Does the singular integral equation

$$\phi(s) = \lambda \int_0^\infty \mathrm{e}^{\mathrm{i}st}\,\phi(t)\,\mathrm{d}t$$

have any nontrivial solutions?

(a) Show that

$$\mathcal{F}\left\{\mathrm{e}^{-a^2t^2}\right\} = \frac{1}{a\sqrt{2}}\,\mathrm{e}^{-s^2/4a^2}.$$

You will need to use the well-known result that

$$\int_{-\infty}^{+\infty} \mathrm{e}^{-x^2}\,\mathrm{d}x = \sqrt{\pi}$$

and the Residue Theorem with an appropriately chosen contour. What can you conclude from the choice $a = 1/\sqrt{2}$?

(b) Show that

$$\mathcal{F}\left\{\frac{1}{|t|^{1/2}}\right\} = \frac{1}{|s|^{1/2}}.$$

Hint: Refer to Eq. (7.12). This exercise shows that the absolute integrability condition is sufficient for a Fourier Transform to exist, but that it is not necessary.

(c) Given that

$$\mathcal{F}\left\{\frac{1}{|t|^{1/4}}\right\} = \frac{\sqrt{\frac{2}{\pi}}\,\Gamma(\tfrac{3}{4})\,\sin(\tfrac{\pi}{8})}{|s|^{3/4}}$$

and

$$\mathcal{F}\left\{\frac{1}{|t|^{3/4}}\right\} = \frac{\sqrt{\frac{2}{\pi}}\,\Gamma(\tfrac{1}{4})\,\cos(\tfrac{\pi}{8})}{|s|^{1/4}},$$

show that

$$\mathcal{F}\left\{\frac{1}{|t|^{1/4}} + \frac{k}{|t|^{3/4}}\right\} = \frac{1}{|s|^{1/4}} + \frac{k}{|s|^{3/4}},$$

where $k = \sqrt{\frac{2}{\pi}}\,\Gamma(\frac{3}{4})\,\sin(\frac{\pi}{8})$. Compare the transforms in this exercise to Eqs. (7.13) and (7.14).

8. The Gaussian kernel with parameter σ is defined by

$$G(s,t;\sigma) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{(s-t)^2}{2\sigma^2}\right\}$$

and the Gaussian Transform is defined in terms of it by

$$\mathcal{G}_\sigma\{f(t)\} = \int_{-\infty}^{+\infty} G(s,t;\sigma)\,f(t)\,\mathrm{d}t.$$

(a) Use *Mathematica* to compute $\mathcal{G}_1\{t^n\}$ for $n = 0, 1, \ldots, 6$.
(b) Solve the singular integral equation $s^6 = \mathcal{G}_1\{p(t)\}$.
(c) Let $\mathcal{P}_N$ denote the vector space consisting of all polynomials of degree less than or equal to N. Show that $\mathcal{G}_1\{\mathcal{P}_N\} = \mathcal{P}_N$.
(d) Does the integral equation $s^n = \mathcal{G}_1\{p(t)\}$ always have a polynomial solution for all integers $n \geq 0$?
(e) Show that

$$\mathrm{e}^s = \frac{1}{\sqrt{e}}\,\mathcal{G}_1\{\mathrm{e}^t\}.$$

(f) Determine λ_σ so that

$$\mathrm{e}^s = \lambda_\sigma\,\mathcal{G}_\sigma\{\mathrm{e}^t\}.$$

Conclude that if $0 < \lambda < 1$, then λ is an eigenvalue of the Gaussian operator.

9. Solve the singular integral equation

$$\phi(x) = f(x) + \frac{1}{2} \int_{-\infty}^{+\infty} \mathrm{e}^{-|x-t|}\, \phi(t)\, \mathrm{d}t.$$

You will need to know that

$$\mathcal{F}\left\{ -\sqrt{\frac{\pi}{2}}\, |t| \right\} = \frac{1}{s^2}.$$

10. Solve the Volterra singular integral equation

$$\phi(x) = \mathrm{e}^x + 2 \int_{x}^{\infty} \mathrm{e}^{-3(x-t)}\, \phi(t)\, \mathrm{d}t.$$

11. Solve the Volterra singular integral equation

$$\phi(x) = \mathrm{e}^x + \int_{-\infty}^{x} \mathrm{e}^{-3(x-t)}\, (x-t)\, \phi(t)\, \mathrm{d}t.$$

12. Solve the singular integral equation

$$\phi(x) = \frac{1}{x} + \lambda \int_{2}^{+\infty} \frac{1}{x^2 t^3}\, \phi(t)\, \mathrm{d}t.$$

Is your solution valid for all λ?

7.3 Equations with Singular Kernels

A *singular kernel* is a kernel that has an infinite discontinuity in the interior of the interval of integration or at a boundary point of it.

7.3.1 Abel's Equation

Abel's equation is the nonsingular integral equation of the first kind of the form

$$f(x) = \int_{0}^{x} \frac{1}{(x-t)^{\alpha}}\, \phi(t)\, \mathrm{d}t, \tag{7.16}$$

where $0 < \alpha < 1$.

Abel showed that solutions to this equation could be found by employing infinite series. Let

$$h(x) = x^{\alpha}\left(\sum_{n=0}^{\infty} a_n x^n\right) = \sum_{n=0}^{\infty} a_n x^{n+\alpha},$$

where $a_0 \neq 0$ and the series $h(x)/x^{\alpha}$ is assumed to have a positive radius of convergence. If we desire a solution of the form

$$\phi(x) = h'(x) = \sum_{n=0}^{\infty} (n+\alpha) a_n x^{n+\alpha-1} = \alpha a_0 x^{\alpha-1} + \cdots$$

then the right-hand side of the integral equation becomes

$$\begin{aligned}\int_0^x \frac{1}{(x-t)^{\alpha}}\,\phi(t)\,\mathrm{d}t &= \sum_{n=0}^{\infty} (n+\alpha)\, a_n \left(\int_0^x (x-t)^{-\alpha} t^{n+\alpha-1}\,\mathrm{d}t\right) \\ &= \sum_{n=0}^{\infty} \frac{1}{n!}\,(n+\alpha)\,\Gamma(n+\alpha)\,\Gamma(1-\alpha)\,a_n x^n \\ &= \frac{\pi}{\sin(\alpha\pi)} \sum_{n=0}^{\infty} \frac{1}{n!}\,\alpha(\alpha+1)\cdots(\alpha+n)\,a_n x^n.\end{aligned}$$

Term-by-term integration here is permissible within the interval of convergence. The integrals here were evaluated and simplified in terms of the beta function and its properties.

On the other hand, if the known inhomogeneous term $f(x)$ has the Maclaurin expansion

$$f(x) = \sum_{n=0}^{\infty} b_n x^n$$

with an assumed positive radius of convergence, where $b_0 \neq 0$, then a comparison of the coefficients in these two series shows that the coefficients of the solution $\phi(x)$ can be written in terms of the coefficients of $f(x)$. Specifically, we have

$$a_n = \frac{\sin(\alpha\pi)}{\pi}\,\frac{n!}{\alpha(\alpha+1)\cdots(\alpha+n)}\,b_n,$$

from which we derive the representation

$$\phi(x) = \frac{\sin(\alpha\pi)}{\pi} \sum_{n=0}^{\infty} \frac{n!}{\alpha(\alpha+1)\cdots(\alpha+n-1)}\,b_n x^{n+\alpha-1}.$$

But since

$$\int_0^x \frac{1}{(x-t)^{1-\alpha}} f(t)\,\mathrm{d}t = \sum_{n=0}^{\infty} b_n \left(\int_0^x (x-t)^{\alpha-1} t^n \,\mathrm{d}t \right)$$
$$= \sum_{n=0}^{\infty} \frac{n!}{\alpha(\alpha+1)\cdots(\alpha+n)} b_n x^{n+\alpha},$$

it follows from these last two equations that the solution $\phi(x)$ has the integral representation

$$\phi(x) = \frac{\sin(\alpha\pi)}{\pi} \frac{\mathrm{d}}{\mathrm{d}x} \left(\int_0^x \frac{1}{(x-t)^{1-\alpha}} f(t)\,\mathrm{d}t \right).$$

It is also possible to solve Abel's equation with the Laplace transform. If $F(s) = \mathcal{L}\{f(x)\}$ and $\Phi(s) = \mathcal{L}\{\phi(x)\}$, then we have the transformed equation

$$F(s) = \frac{\Gamma(1-\alpha)}{s^{1-\alpha}} \Phi(s),$$

which can be rearranged in the form

$$\frac{\Phi(s)}{s} = \frac{s^{-\alpha}\Gamma(\alpha)}{\Gamma(1-\alpha)\Gamma(\alpha)} F(s) = \frac{\sin(\alpha\pi)}{\pi} \frac{\Gamma(\alpha)}{s^{\alpha}} F(s).$$

After inversion, we obtain

$$\mathcal{L}\left\{ \int_0^x \phi(t)\,\mathrm{d}t \right\} = \frac{\sin(\alpha\pi)}{\pi} \mathcal{L}\left\{x^{\alpha-1}\right\} \mathcal{L}\{f(x)\}$$
$$= \frac{\sin(\alpha\pi)}{\pi} \mathcal{L}\left\{ \int_0^x \frac{1}{(x-t)^{1-\alpha}} f(t)\,\mathrm{d}t \right\},$$

from which we conclude that

$$\phi(x) = \frac{\sin(\alpha\pi)}{\pi} \frac{\mathrm{d}}{\mathrm{d}x} \left(\int_0^x \frac{1}{(x-t)^{1-\alpha}} f(t)\,\mathrm{d}t \right).$$

A third approach to Abel's equation is possible. If we replace x by s in Eq. (7.16), multiply by $\mathrm{d}s/(x-s)^{1-\alpha}$, and then integrate with respect to s, we obtain

$$\int_{s=0}^{x} \frac{1}{(x-s)^{1-\alpha}} f(s)\,\mathrm{d}s = \int_{s=0}^{x} \left(\int_{t=0}^{s} \frac{1}{(x-s)^{1-\alpha}(s-t)^{\alpha}} \phi(t)\,\mathrm{d}t \right) \mathrm{d}s$$
$$= \int_{t=0}^{x} \left(\int_{s=t}^{x} \frac{1}{(x-s)^{1-\alpha}(s-t)^{\alpha}} \mathrm{d}s \right) \phi(t)\,\mathrm{d}t,$$

after interchanging the order of integration. (The student should justify this step after considering appropriate assumptions on $\phi(t)$ such as continuity.) The inner integral on the right-hand side can be evaluated as a beta integral with a simple substitution. If we let $u = (s-t)/(x-t)$, so that $\mathrm{d}s = (x-t)\,\mathrm{d}u$, then

$$\int_{s=t}^{x} \frac{1}{(x-s)^{1-\alpha}(s-t)^{\alpha}} \mathrm{d}s = \int_{0}^{1} u^{-\alpha}(1-u)^{\alpha-1}\,\mathrm{d}u = \frac{\pi}{\sin(\alpha\pi)}.$$

Hence, we have

$$\int_{0}^{x} \phi(t)\,\mathrm{d}t = \frac{\sin(\alpha\pi)}{\pi} \int_{0}^{x} \frac{1}{(x-s)^{1-\alpha}} f(s)\,\mathrm{d}s,$$

from which we derive

$$\phi(x) = \frac{\sin(\alpha\pi)}{\pi} \frac{\mathrm{d}}{\mathrm{d}x} \left(\int_{0}^{x} \frac{1}{(x-s)^{1-\alpha}} f(s)\,\mathrm{d}s \right).$$

If $f(s)$ is differentiable, then the solution to Abel's equation has another formulation. A simple integration by parts shows that

$$\int_{0}^{x} \frac{1}{(x-s)^{1-\alpha}} f(s)\,\mathrm{d}s = \frac{1}{\alpha} x^{\alpha} f(0) + \frac{1}{\alpha} \int_{0}^{x} (x-s)^{\alpha} f'(s)\,\mathrm{d}s,$$

if $f(0)$ is defined. It now follows that

$$\phi(t) = \frac{\sin(\alpha\pi)}{\pi} \left(\frac{f(0)}{x^{1-\alpha}} + \int_{0}^{x} \frac{1}{(x-s)^{1-\alpha}} f'(s)\,\mathrm{d}s \right).$$

Illustrative Examples

- *Example 1*: Abel originally considered Eq. (7.16) with $\alpha = \frac{1}{2}$ in the course of investigating the tautochrone problem whose solution is widely available in the literature.[4]

  Although Eq. (7.16) is sometimes referred to as the *generalized Abel equation*, even more general versions of it exist. For example, if $B(x,t)$ is bounded and continuous in the triangle $T(0,1)$ with $B(x,x) \neq 0$, then the equation

$$f(x) = \int_0^x \frac{B(x,t)}{(x-t)^\alpha}\,\phi(t)\,\mathrm{d}t$$

  has also been considered. It is solved by converting it to an equivalent Volterra integral equation of the first kind as follows. If we replace x by s, multiply by $1/(x-s)^{1-\alpha}$, and then integrate the result, we obtain after some simplification, the equation

$$\tilde{f}_\alpha(x) = \int_0^x J(x,t)\,\phi(t)\,\mathrm{d}t,$$

  where

$$J(x,t) = \int_t^x \frac{B(s,t)}{(x-s)^{1-\alpha}(s-t)^\alpha}\,\mathrm{d}s = \int_0^1 \frac{B(t+(x-t)u,t)}{(1-u)^{1-\alpha}u^\alpha}\,\mathrm{d}u$$

  and

$$\tilde{f}_\alpha(x) = \int_0^x \frac{1}{(x-s)^{1-\alpha}}\,f(s)\,\mathrm{d}s$$

  are continuously differentiable with respect to x. The converted equation can be differentiated, if necessary, to obtain a Volterra integral equation of the second kind.

  Another type of generalization is possible. If $\gamma\colon [0,1] \to \mathbb{R}$ has a positive and continuous derivative, then the singular integral equation

$$f(x) = \int_0^x \frac{1}{(\gamma(x)-\gamma(t))^\alpha}\,\phi(t)\,\mathrm{d}t$$

[4]For a discussion of this problem, see Andrews, L.C. and Shivamoggi, B.K., *Integral Transforms for Engineers and Applied Mathematicians*, MacMillan Publishing Company, New York, 1988, pp. 240–242.

has the solution

$$\phi(x) = \frac{\sin(\alpha\pi)}{\pi} \frac{\mathrm{d}}{\mathrm{d}x} \left(\int_0^x \frac{\gamma'(t)}{(\gamma(x) - \gamma(t))^{1-\alpha}} f(t)\,\mathrm{d}t \right). \tag{7.17}$$

- *Example 2*: Abel's integral equation of the second kind is given by

$$\phi(x) = f(x) + \lambda \int_0^x \frac{1}{\sqrt{x-t}} \phi(t)\,\mathrm{d}t.$$

Assume that $\phi(x)$ is continuous at $x = 0$. If we replace x by s, multiply by $\mathrm{d}s/\sqrt{x-s}$, and then integrate with respect to s, we obtain

$$\int_0^x \frac{1}{\sqrt{x-s}} \phi(s)\,\mathrm{d}s = \int_0^x \frac{1}{\sqrt{x-s}} f(s)\,\mathrm{d}s + \lambda \int_0^x \int_0^s \frac{1}{\sqrt{x-s}\sqrt{s-t}} \phi(t)\,\mathrm{d}t\,\mathrm{d}s.$$

Next, if we interchange the order of integration in the double integral and then multiply by λ, we obtain

$$\lambda \int_0^x \frac{1}{\sqrt{x-s}} \phi(s)\,\mathrm{d}s = \lambda \int_0^x \frac{1}{\sqrt{x-s}} f(s)\,\mathrm{d}s + \lambda^2 \pi \int_0^x \phi(t)\,\mathrm{d}t.$$

After subtracting Abel's equation from the last equation, we obtain

$$\phi(x) = g(x) + \lambda^2 \pi \int_0^x \phi(t)\,\mathrm{d}t,$$

where we have set

$$g(x) = f(x) + \lambda \int_0^x \frac{1}{\sqrt{x-s}} f(s)\,\mathrm{d}s.$$

Upon differentiation, we find that $\phi(x)$ must satisfy the first-order linear differential equation

$$\phi'(x) - \lambda^2 \pi \phi(x) = g'(x).$$

After integrating this differential equation, we obtain

$$\mathrm{e}^{-\lambda^2 \pi x} \phi(x) - \phi(0) = \int_0^x \mathrm{e}^{-\lambda^2 \pi t} g'(t)\,\mathrm{d}t.$$

Integration by parts, when applied to the integral on the right, yields

$$e^{-\lambda^2\pi x}\phi(x) - \phi(0) = e^{-\lambda^2\pi x} g(x) - g(0) + \lambda^2\pi \int_0^x e^{-\lambda^2\pi t} g(t)\,dt.$$

But since $\phi(0) = g(0)$, we finally have

$$\phi(x) = g(x) + \lambda^2\pi \int_0^x e^{\lambda^2\pi(x-t)} g(t)\,dt.$$

7.3.2 Equations with Cauchy and Hilbert Kernels

In this subsection, we consider integral equations with singular kernels of great theoretical and practical importance. It is important to realize at the outset that many singular integral equations involve integrals that exist only in the sense of Cauchy principal value and, in addition, that the solutions to those equations will accordingly exist only in the sense of Cauchy principal value as well.

The *Cauchy kernel* is given by

$$C(x,t) = \frac{1}{x-t},$$

and the *Hilbert kernel* is given by

$$H(x,t) = \cot\left(\frac{x-t}{2}\right).$$

These two kernels are analytically related since

$$\frac{1}{x-t} = \frac{1}{2}\cot\left(\frac{x-t}{2}\right) + \sum_{n=1}^{\infty} \frac{(-1)^{n-1} B_{2n}}{(2n)!}(x-t)^{2n-1}, \tag{7.18}$$

where the constants B_{2n} are the Bernoulli numbers.

Before considering singular integral equations with either of these kernels, it is necessary to show that the singular integrals themselves exist in the sense of the Cauchy principal value under reasonable conditions. One very common and simple condition is the following.

The function $f(x)$ is said to satisfy a *Lipschitz condition with exponent* α if

$$|f(x_1) - f(x_2)| \le C\,|x_1 - x_2|^{\alpha}$$

for all $x_1, x_2 \in [a,b]$, where C is a positive constant and $0 < \alpha \le 1$. Then

$$\left|\frac{f(t) - f(x)}{t-x}\right| \le \frac{C}{|t-x|^{1-\alpha}}$$

for all $t, x \in [a,b]$. Now let $x \in [a,b]$ and consider the identity

$$\int_a^b \frac{f(t)}{t-x}\,\mathrm{d}t = \int_a^b \frac{f(t)-f(x)}{t-x}\,\mathrm{d}t + f(x)\int_a^b \frac{1}{t-x}\,\mathrm{d}t.$$

The first integral on the right exists, albeit as an improper integral, and the second integral on the right exists only in the sense of Cauchy principal value. Hence,

$$PV\int_a^b \frac{f(t)}{t-x}\,\mathrm{d}t = \int_a^b \frac{f(t)-f(x)}{t-x}\,\mathrm{d}t + f(x)\ln\left(\frac{b-x}{x-a}\right).$$

A similar observation can be made for singular integrals with Hilbert kernels.

A theorem due to Privalov[5] demonstrates that higher iterates also exist in this sense. Let

$$f_1(x) = PV\int_a^b \frac{f(t)}{t-x}\,\mathrm{d}t$$

and

$$f_{n+1}(x) = PV\int_a^b \frac{f_n(t)}{t-x}\,\mathrm{d}t$$

for $n \geq 1$. If $\alpha < 1$, then $f_1(x)$ satisfies a Lipschitz condition with exponent α on the interval $[c,d]$ where $a < c < d < b$. If $\alpha = 1$, then $f_1(x)$ satisfies a Lipschitz condition for arbitrary $\beta < 1$ on $[c,d]$. It follows that $f_2(x)$ exists. A short corollary to Privalov's theorem shows that $f_n(x)$ exists for all $n \geq 1$ under these same conditions.

Without loss of generality, we can assume that $a = -1$ and that $b = +1$, since the change of limits can be effected by an elementary change of variables. Another common choice of limits is $a = 0$ and $b = +1$.

If the interval of integration is infinite, then we must write

$$PV\int_{-\infty}^{+\infty} \frac{f(t)}{t-x}\,\mathrm{d}t = \int_{-\infty}^{a} \frac{f(t)}{t-x}\,\mathrm{d}t + PV\int_a^b \frac{f(t)}{t-x}\,\mathrm{d}t + \int_b^{+\infty} \frac{f(t)}{t-x}\,\mathrm{d}t,$$

where $x \in (a,b)$. If $f(x)$ is square integrable on the interval $(-\infty,+\infty)$, then the first integral on the right exists as an application of the Cauchy–Schwarz Inequality, since

$$\left|\int_{-\infty}^{a} \frac{f(t)}{t-x}\,\mathrm{d}t\right|^2 \leq \left(\int_{-\infty}^{a} f^2(t)\,\mathrm{d}t\right)\left(\int_{-\infty}^{a}\left(\frac{1}{t-x}\right)^2\mathrm{d}t\right) < +\infty.$$

A similar comment can be made for the third integral on the right.

[5]For a proof of this theorem, see Privalov, I.I., *Introduction to the Theory of Functions of a Complex Variable*, 8th edn. State Technical Publishers, 1948.

The singular integrals have been defined here for functions of a real variable x on an interval. In more advanced texts, it is shown that analogous integrals can also be defined for functions of a complex variable z on smooth contours in the complex plane. Interesting integral equations can be posed involving contour integrals, although we shall not do so here.

Three Special Integral Transforms

- *1. The Cauchy Kernel (finite interval):* If $f(x)$ satisfies a Lipschitz condition on the interval $(-1,+1)$, then the *Finite Hilbert Transform*

$$\mathbf{h}\colon f(x) \to \frac{1}{\pi} PV \int_{-1}^{+1} \frac{1}{t-x} f(t)\,\mathrm{d}t = h(x)$$

exists in the sense of the Cauchy principal value. This transformation has many interesting properties. One of the most useful is the *convolution theorem*, which states that

$$\mathbf{h}[\phi\,\mathbf{h}\psi + \psi\mathbf{h}\phi] = \mathbf{h}\phi\,\mathbf{h}\psi - \phi\,\psi. \tag{7.19}$$

In the exercises, it is shown that $\mathbf{h}\left[(1-x^2)^{-1/2}\right] = 0$, from which it follows that $\mathbf{h}\left[(1-x^2)^{1/2}\right] = -x$.

Now consider the singular integral equation of the first kind

$$f(x) = \frac{1}{\pi} PV \int_{-1}^{+1} \frac{1}{t-x} \phi(t)\,\mathrm{d}t,$$

known as the *airfoil equation*, which can be more simply written in the form $f = \mathbf{h}\phi$. It is easy to determine the form that any solution to this equation must assume. If we set $\psi(x) = (1-x^2)^{1/2}$ in the convolution theorem, then we readily obtain the relation

$$\mathbf{h}\left[-x\phi(x) + (1-x^2)^{1/2} f(x)\right] = -x f(x) - (1-x^2)^{1/2}\phi(x).$$

In order to isolate $\phi(x)$, observe that

$$\mathbf{h}[x\phi(x)] = \frac{1}{\pi} PV \int_{-1}^{+1} \frac{1}{t-x} t\,\phi(t)\,\mathrm{d}t$$

$$= \frac{1}{\pi} PV \int_{-1}^{+1} \frac{(t-x)+x}{t-x} \phi(t)\,\mathrm{d}t$$

$$= \frac{1}{\pi} \int_{-1}^{+1} \phi(t)\,\mathrm{d}t + \frac{x}{\pi} PV \int_{-1}^{+1} \frac{1}{t-x}\,\phi(t)\,\mathrm{d}t$$

$$= \frac{1}{\pi} \int_{-1}^{+1} \phi(t)\,\mathrm{d}t + x f(x).$$

After substituting the result of this computation into the equation above it, we obtain

$$(1-x^2)^{1/2}\,\phi(x) = -\mathbf{h}\left[(1-x^2)^{1/2} f(x)\right] + \frac{1}{\pi} \int_{-1}^{+1} \phi(t)\,\mathrm{d}t,$$

or, in terms of an explicit integral,

$$\phi(x) = -\frac{1}{\pi} PV \int_{-1}^{+1} \sqrt{\frac{1-t^2}{1-x^2}}\,\frac{1}{t-x}\,f(t)\,\mathrm{d}t + \frac{k}{\sqrt{1-x^2}},$$

where we have set

$$k = \frac{1}{\pi} \int_{-1}^{+1} \phi(t)\,\mathrm{d}t.$$

- *2. The Cauchy Kernel (infinite interval):* If $f(x)$ is square integrable and satisfies a Lipschitz condition on the interval $(-\infty, +\infty)$, then the *Hilbert Transform*

$$\mathbf{H}\colon\ f(x) \to \frac{1}{\pi} PV \int_{-\infty}^{+\infty} \frac{1}{t-x} f(t)\,\mathrm{d}t = h(x)$$

exists in the sense of the Cauchy principal value. Furthermore, $h(x)$ is also square integrable and satisfies the same Lipschitz condition as $f(x)$, so that its Hilbert Transform

$$\mathbf{H}\colon\ h \to \frac{1}{\pi} PV \int_{-\infty}^{+\infty} \frac{1}{t-x} h(t)\,\mathrm{d}t$$

also exists. One of the most surprising and important properties of the Hilbert Transform, derived from the Fourier Integral Theorem, is the fact that if $h = \mathbf{H}f$, then $\mathbf{H}h = -f$, from which it follows that $\mathbf{HH}f = -f$.

Now consider the singular integral equation of the second kind

$$\phi(x) = f(x) + \lambda \int_{-\infty}^{+\infty} \frac{1}{t-x}\,\phi(t)\,\mathrm{d}t,$$

where we assume that $f(x)$ is square integrable and satisfies a Lipschitz condition on the interval $(-\infty,+\infty)$. In operator form, this equation can be alternately written as

$$\phi = f + \lambda\pi\mathbf{H}\phi.$$

If we apply $\mathbf{H}$ to this equation, then we obtain

$$\mathbf{H}\phi = \mathbf{H}f + \lambda\pi\mathbf{H}\mathbf{H}\phi \quad \text{or} \quad \mathbf{H}\phi = \mathbf{H}f - \lambda\pi\phi.$$

After eliminating $\mathbf{H}\phi$ from these last two equations, we obtain

$$\phi = \frac{1}{1+\lambda^2\pi^2}\left(f + \lambda\pi\mathbf{H}f\right),$$

or

$$\phi(x) = \frac{1}{1+\lambda^2\pi^2}\left(f + \lambda\, PV\int_{-\infty}^{+\infty}\frac{1}{t-x}f(t)\,\mathrm{d}t\right),$$

where $1+\lambda^2\pi^2 \neq 0$. The solution $\phi(x)$ is square integrable on the interval $(-\infty,+\infty)$.

- *3. The Hilbert Kernel:* Another closely related integral transform is defined in terms of the Hilbert kernel. Let

$$\mathbf{P}\colon f(x) \to \frac{1}{2\pi}PV\int_0^{2\pi}\cot\left(\frac{x-t}{2}\right)f(t)\,\mathrm{d}t = p(x).$$

Given the close analytic relation (7.18) between the Cauchy kernel and the Hilbert kernel, we can assert that $\mathbf{P}f$ exists if $f(x)$ satisfies a Lipschitz condition on the interval $(0,2\pi)$. Note that if k is a constant, then $\mathbf{P}k = 0$, due to the periodicity of the cotangent.

The most important property of this transform, relevant to the purpose of solving singular integral equations with Hilbert kernels, is the observation that if $p = \mathbf{P}f$, then $\mathbf{P}p = \mathbf{P}\mathbf{P}f = -f + A[f]$, where

$$A[f] = \frac{1}{2\pi}\int_0^{2\pi}f(s)\,\mathrm{d}s$$

is the average value of $f(s)$ on the interval $[0,2\pi]$.

Consider the singular integral equation of the first kind

$$\frac{1}{2\pi}PV\int_0^{2\pi}\cot\left(\frac{x-t}{2}\right)\phi(t)\,\mathrm{d}t = f(x),$$

which can be written more simply in operator notation as $\mathbf{P}\phi = f$. If we apply $\mathbf{P}$ to both sides of the integral equation, we obtain $\mathbf{PP}\phi = \mathbf{P}f$. However, since $\mathbf{PP}\phi = -\phi + A[\phi]$, the solution to the integral equation is

$$\phi = -\mathbf{P}f + A[\phi],$$

or, in explicit integral form,

$$\phi(x) = -\frac{1}{2\pi} PV \int_0^{2\pi} \cot\left(\frac{x-t}{2}\right) f(t)\,\mathrm{d}t + \frac{1}{2\pi}\int_0^{2\pi} \phi(s)\,\mathrm{d}s.$$

To show that ϕ satisfies the equation, we apply $\mathbf{P}$ to the solution $\phi = -\mathbf{P}f + A[\phi]$ to obtain

$$f = \mathbf{P}\phi = -\mathbf{PP}f + \mathbf{P}A[\phi] = f(x) - A[f],$$

showing that a necessary and sufficient condition for a solution to exist is that $A[f] = 0$.

Now consider the singular integral equation of the second kind

$$a\,\phi(x) = f(x) + \frac{b}{2\pi} PV \int_0^{2\pi} \cot\left(\frac{x-t}{2}\right) \phi(t)\,\mathrm{d}t,$$

where a and b are constants, possibly complex.

Let

$$\mathbf{L}^{-}\phi = a\,\phi(x) = \phi(x) - \frac{b}{2\pi} PV \int_0^{2\pi} \cot\left(\frac{x-t}{2}\right) \phi(t)\,\mathrm{d}t$$

and

$$\mathbf{L}^{+}\phi = a\,\phi(x) = \phi(x) + \frac{b}{2\pi} PV \int_0^{2\pi} \cot\left(\frac{x-t}{2}\right) \phi(t)\,\mathrm{d}t.$$

The integral equation can be written in the alternate form $\mathbf{L}^{-}\phi = f$. If we apply $\mathbf{L}^{+}$ to both sides of this equation, we obtain

$$\mathbf{L}^{+}\mathbf{L}^{-}\phi = \mathbf{L}^{+}f = f + b\mathbf{P}f.$$

But also,

$$\begin{aligned}
\mathbf{L}^{+}\mathbf{L}^{-}\phi &= \mathbf{L}^{+}\left[a\,\phi - b\mathbf{P}\phi\right] \\
&= a\mathbf{L}^{+}\phi - b\mathbf{L}^{+}\left[\mathbf{P}\phi\right] \\
&= a\left(a\,\phi + b\mathbf{P}\phi\right) - b\left(a\mathbf{P}\phi + b\mathbf{PP}\phi\right) \\
&= a^2\,\phi - b^2\left(-\phi + A[\phi]\right) \\
&= \left(a^2 + b^2\right)\phi - b^2 A[\phi].
\end{aligned}$$

It follows from these last two equations that

$$\left(a^2+b^2\right)\,\phi(x) = a\,f(x) + b\,\mathbf{P}f + \frac{b^2}{2\pi}\int_0^{2\pi}\phi(s)\,\mathrm{d}s.$$

We have shown that the solution $\phi(x)$ to the singular integral equation is the solution to a simple Fredholm equation with the kernel $K(x,s)\equiv 1$. Actually, it can be shown that these two integral equations are equivalent. In order to complete the solution, we must evaluate $A[\phi]$. Upon integrating the last equation, we find that

$$A[\phi] = \frac{1}{2\pi a}A[f].$$

Thus, if $a^2+b^2\neq 0$, then the solution to the integral equation is given by

$$\phi(x) = \frac{a}{a^2+b^2}\,f(x) + \frac{b}{a^2+b^2}\,\mathbf{P}f + \frac{b^2}{a\,(a^2+b^2)}A[f].$$

If $a^2+b^2=0$, then $\mathbf{L}^+f = -b^2A[\phi]$, i.e., $\mathbf{L}^+f$ is constant. In this case, the integral equation does not in general have a solution.

Illustrative Examples

- *Example 1*: Consider the singular integral equation of the first kind

$$f(x) = \frac{1}{1+x^2} = \frac{1}{\pi}\,PV\int_{-\infty}^{+\infty}\frac{1}{t-x}\,\phi(t)\,\mathrm{d}t = \mathbf{H}\phi.$$

Before we attempt to solve this equation, it is necessary to note that $f(x)$ is square integrable and that it satisfies a Lipschitz condition with exponent 1. As a consequence of the Mean Value Theorem, it is easy to show that

$$|f(x_1)-f(x_2)| = \left|f'(c)\right|\,|x_1-x_2| \leq \frac{3\sqrt{3}}{8}\,|x_1-x_2|$$

for any real x_1 and x_2, with $x_1<c<x_2$. It follows that $\mathbf{H}f$ exists.

Since $\mathbf{H}f = \mathbf{H}\mathbf{H}\phi = -\phi$, the solution to the integral equation is

$$\phi(x) = -\mathbf{H}f = -\frac{1}{\pi}\,PV\int_{-\infty}^{+\infty}\frac{1}{t-x}\,\frac{1}{1+t^2}\,\mathrm{d}t.$$

This integral is evaluated by employing the method of partial fractions. Let

$$\begin{aligned} I(t) &= \int \frac{1}{t-x}\,\frac{1}{1+t^2}\,\mathrm{d}t \\ &= \frac{1}{1+x^2}\int\left(\frac{1}{t-x}-\frac{t}{1+t^2}-\frac{x}{1+t^2}\right)\mathrm{d}t \\ &= \frac{1}{1+x^2}\left(\ln|t-x|-\frac{1}{2}\ln\left|1+t^2\right|-x\arctan t\right). \end{aligned}$$

It is easy to see that:

- As $t\to-\infty$, $I(t)\to\frac{\pi x}{2+2x^2}$.
- As $t\to+\infty$, $I(t)\to-\frac{\pi x}{2+2x^2}$.
- At $t=x\pm\varepsilon$, $I(x\pm\varepsilon)=\frac{1}{1+x^2}[\ln|\frac{\pm\varepsilon}{\sqrt{1+(x\pm\varepsilon)^2}}|-x\arctan|x\pm\varepsilon|]$.

We conclude from these values that

$$\begin{aligned} \phi(x) &= -\frac{1}{\pi}\,\frac{1}{1+x^2}\lim_{\varepsilon\downarrow0}\left(\int_{-\infty}^{x-\varepsilon}+\int_{x+\varepsilon}^{+\infty}\right)\frac{1}{t-x}\,\frac{1}{1+t^2}\,\mathrm{d}t \\ &= \frac{x}{1+x^2}. \end{aligned}$$

Finally, we note that $\phi(x)$ is square integrable and satisfies a Lipschitz condition.

- *Example 2*: Compute

$$\mathbf{P}[\sin x]=\frac{1}{2\pi}\,PV\int_0^{2\pi}\cot\left(\frac{x-t}{2}\right)\sin t\,\mathrm{d}t.$$

With the linear substitution $u=(x-t)/2$ and elementary trigonometric identities, the indefinite integral is evaluated to be

$$\begin{aligned} J(t) &= \int\cot\left(\frac{x-t}{2}\right)\sin t\,\mathrm{d}t \\ &= -t\cos x-\sin t-2\ln\left|\sin\left(\frac{x-t}{2}\right)\right|+C. \end{aligned}$$

(Terms involving x only are absorbed into C.)

It is easy to see that:

- $J(0)=-2\ln|\sin(\frac{x}{2})|+C$.
- $J(2\pi)=-2\pi\cos x-2\ln|\sin(\frac{x}{2})|+C$.
- $J(x\pm\varepsilon)=-(x\pm\varepsilon)\cos x-\sin(x\pm\varepsilon)-2\ln|\sin(\mp\varepsilon)|+C$.

We conclude from these values that

$$\mathbf{P}[\sin x] = \frac{1}{2\pi} \lim_{\varepsilon \downarrow 0} \left(\int_0^{x-\varepsilon} + \int_{x+\varepsilon}^{2\pi} \right) \cot\left(\frac{x-t}{2}\right) \sin t \, \mathrm{d}t = -\cos x.$$

- *Example 3*: The singular integral equation

$$\int_{-1}^{+1} \ln|x-t| \, \phi(t) \, \mathrm{d}t = 1$$

with a logarithmic kernel can be solved by employing Fourier series.

Since $-1 < x,t < +1$, the change of variables $x = \cos y$ and $t = \cos u$ yields the reformulation

$$\int_0^{\pi} \ln|\cos y - \cos u| \, \psi(u) \, \mathrm{d}u = 1$$

where $\psi(u) = \phi(\cos u) \sin u$.

The well-known Fourier series

$$\frac{1}{2} \ln(1-\cos\theta) = -\frac{1}{2} \ln 2 - \sum_{n=1}^{\infty} \frac{1}{n} \cos(n\theta) \quad (0 < \theta < 2\pi)$$

can be adapted to suit our purpose by using only elementary trigonometric identities. Indeed, if we let $L(y,u) = \ln|\cos y - \cos u|$, then we have

$$\begin{aligned}
L(y,u) &= \ln\left|2 \sin\left(\frac{y+u}{2}\right) \sin\left(\frac{u-y}{2}\right)\right| \\
&= \frac{1}{2} \ln\left(2\sin^2\left(\frac{y+u}{2}\right)\right) + \frac{1}{2} \ln\left(2 \sin^2\left(\frac{u-y}{2}\right)\right) \\
&= \frac{1}{2} \ln\left(1-\cos\left(\frac{y+u}{2}\right)\right) + \frac{1}{2} \ln\left(1-\cos\left(\frac{u-y}{2}\right)\right) \\
&= -\frac{1}{2} \ln 2 - \sum_{n=1}^{\infty} \frac{1}{n} \cos\left(n\left(\frac{y+u}{2}\right)\right) \\
&\quad -\frac{1}{2} \ln 2 - \sum_{n=1}^{\infty} \frac{1}{n} \cos\left(n\left(\frac{u-y}{2}\right)\right) \\
&= -\ln 2 - \sum_{n=1}^{\infty} \frac{1}{n} \left[\cos\left(n\left(\frac{y+u}{2}\right)\right) + \cos\left(n\left(\frac{u-y}{2}\right)\right)\right] \\
&= -\ln 2 - 2 \sum_{n=1}^{\infty} \frac{1}{n} \cos(ny)\cos(nu).
\end{aligned}$$

Assuming that the solution $\phi(t)$ is continuous with a piecewise continuous derivative on $(-1,+1)$, the composed function $\psi(u)$ is also continuous with a piecewise continuous deivative on $(0,\pi)$. Furthermore, the even extension of $\psi(u)$ to $(-\pi,+\pi)$ will have a convergent Fourier expansion of the form

$$\psi(u) = \sum_{n=0}^{\infty} b_n \cos(nu)$$

which converges to $\psi(u)$ on the interval $(0,\pi)$.

If we substitute these two Fourier series into the reformulated integral equation and evaluate the resulting integral, then we obtain

$$-\pi b_0 \ln 2 - \sum_{n=1}^{\infty} \frac{\pi b_n}{n} \cos(ny) = 1$$

for all $y \in (0,\pi)$ due to the orthogonality properties of the cosine function. It follows that $-\pi b_0 \ln 2 = 1$ and $b_n = 0$ for all $n \geq 1$. Consequently,

$$\psi(u) = b_0 = -\frac{1}{\pi \ln 2}$$

or, equivalently,

$$\phi(\cos u) \sin u = \phi(\cos u) \sqrt{1-\cos^2 u} = -\frac{1}{\pi \ln 2}.$$

Finally, the solution to the integral equation is given by

$$\phi(x) = -\frac{1}{\pi \ln 2} \frac{1}{\sqrt{1-x^2}}.$$

If we substitute this solution back into the given integral equation, then we obtain as a corollary the evaluated definite integral

$$\int_{-1}^{+1} \ln|x-t| \frac{1}{\sqrt{1-t^2}} \, \mathrm{d}t = -\pi \ln 2. \tag{7.20}$$

The general singular integral equation of the first kind

$$\int_{-1}^{+1} \ln|x-t| \, \phi(t) \, \mathrm{d}t = f(x) \tag{7.21}$$

with a logarithmic kernel can now be solved.

If we differentiate (justify!) this integral equation, then we obtain

$$\frac{1}{\pi} PV \int_{-1}^{+1} \frac{1}{t-x} \phi(t) = -\frac{1}{\pi} f'(x),$$

or, equivalently, $\mathbf{h}\phi = -f'(x)/\pi$, where $\mathbf{h}$ is the Finite Hilbert Transform.

We recognize this equation as a singular integral equation with a Cauchy kernel (the airfoil equation) which was solved in the exposition above. Its solution was determined there to be

$$\phi(x) = +\frac{1}{\pi^2} PV \int_{-1}^{+1} \sqrt{\frac{1-t^2}{1-x^2}} \frac{1}{t-x} f'(t)\,\mathrm{d}t + \frac{k}{\sqrt{1-x^2}},$$

where we have set

$$k = \frac{1}{\pi} \int_{-1}^{+1} \phi(t)\,\mathrm{d}t.$$

It remains to evaluate k. If we replace x with u in Eq. (7.21), multiply by $1/\sqrt{1-u^2}$, integrate with respect to u, and then interchange the order of integration, we obtain

$$\int_{t=-1}^{+1} \left(\int_{u=-1}^{+1} \ln|u-t| \frac{1}{\sqrt{1-u^2}}\,\mathrm{d}u \right) \phi(t)\,\mathrm{d}t = \int_{-1}^{+1} \frac{1}{\sqrt{1-u^2}} f(u)\,\mathrm{d}u.$$

Given the value (7.20), we see that

$$k = -\frac{1}{\pi^2 \ln 2} \int_{-1}^{+1} \frac{1}{\sqrt{1-u^2}} f(u)\,\mathrm{d}u.$$

The final form of the solution

$$\phi(x) = \frac{1}{\pi^2 \sqrt{1-x^2}} \left[PV \int_{-1}^{+1} \frac{\sqrt{1-t^2}}{t-x} f'(t)\,\mathrm{d}t - \frac{1}{\ln 2} \int_{-1}^{+1} \frac{1}{\sqrt{1-u^2}} f(u)\,\mathrm{d}u \right]$$

is called Carleman's formula.

It is clear from this form of the solution that suitable assumptions must be made a priori on $f'(x)$ and $f(x)$ for these integrals to exist. For example, if $f'(x)$ satisfies a Lipschitz condition on the interval $(-1,+1)$, then the first integral will exist in the sense of the Cauchy principal value. The second integral must exist as an improper integral since $1/\sqrt{1-t^2}$ becomes unbounded as $t \to \pm 1$. It will exist, for example, if $f(x)$ is bounded and integrable.

Sect. 7.3 Exercises

1. Let $0 < \alpha < 1$ and let n be a positive integer. The solution to Abel's equation

$$x^n = \int_0^x \frac{1}{(x-t)^\alpha}\,\phi(t)\,\mathrm{d}t$$

has the form $\phi(t) = cx^d$. Determine c and d.

If $P_N(x)$ is a polynomial of degree N, does the integral equation

$$P_N(x) = \int_0^x \frac{1}{(x-t)^\alpha}\,\phi(t)\,\mathrm{d}t$$

always have a solution?

2. For real values of α and β, when does the integral equation

$$x^\beta = \int_0^x \frac{1}{(x-t)^\alpha}\,\phi(t)\,\mathrm{d}t$$

have a solution?

3. Derive the solution (7.17) by following the method outlined in the text.
4. Abel's equation is easily converted to the form

$$\tilde{f}(y) = \int_y^1 \frac{1}{(u-y)^\alpha}\,\tilde{\phi}(u)\,\mathrm{d}u$$

by setting $x = 1-y, t = 1-u$, $\tilde{f}(y) = f(1-y)$, and $\tilde{\phi}(u) = \phi(1-u)$.

(a) Determine the general form of the solution $\tilde{\phi}(y)$ to this equation.
(b) Solve the equation

$$1 - y = \int_y^1 \frac{1}{(u-y)^\alpha}\,\tilde{\phi}(u)\,\mathrm{d}u.$$

5. Solve Abel's singular integral equation

$$f(x) = \int_x^{+\infty} \frac{1}{\sqrt{x-t}}\,\phi(t)\,\mathrm{d}t.$$

Hint: Replace x by s, multiply by $1/\sqrt{x-s}$, and then integrate from x to $+\infty$.

6. Evaluate the Finite Hilbert Transform

$$\mathbf{h}\left[(1-x^2)^{-1/2}\right] = \frac{1}{\pi}\,PV\int_{-1}^{+1}(1-t^2)^{-1/2}\,\frac{1}{t-x}\,\mathrm{d}t.$$

(a) Explain why $(1-x^2)^{-1/2}$ *does not* satisfy a Lipschitz condition on the interval $(-1,+1)$.
(b) Let $-1 < a < x < b < +1$. Use the Mean Value Theorem to explain why $(1-x^2)^{-1/2}$ *does* satisfy a Lipschitz condition on the interval $[a,b]$.
(c) Explain why $\mathbf{h}\left[(1-x^2)^{-1/2}\right]$ exists, regardless of part (a).
(d) Use the substitution

$$t = \frac{1-s^2}{1+s^2}, \qquad \mathrm{d}t = -\frac{4s}{(1+s^2)^2}$$

to show that

$$\mathbf{h}\left[(1-x^2)^{-1/2}\right] = \frac{2}{\pi} PV \int_0^{+\infty} \frac{1}{(1-x)-(1+x)s^2}\,\mathrm{d}s.$$

Note that t is a decreasing function of s, so that $t = +1$ corresponds to $s = 0$ and that $t = -1$ corresponds to $s = +\infty$. Also, since $0 < 1 \pm x < 2$, there is an asymptote at $\tilde{s} = \sqrt{(1-x)/(1+x)}$, with $0 < \tilde{s} < +\infty$.
(e) Use the antiderivative formula

$$\int \frac{1}{a^2 - b^2 s^2}\,\mathrm{d}s = \frac{1}{2ab} \ln\left|\frac{a+bs}{a-bs}\right| + C$$

to evaluate the integral in part (d) on the set $[0, \tilde{s}-\varepsilon] \cup [\tilde{s}+\varepsilon, +\infty)$.
(f) Conclude that $\mathbf{h}\left[c\,(1-x^2)^{-1/2}\right] = 0$, where c is a constant. Are there any other nontrivial functions ψ for which $\mathbf{h}\psi = 0$?
(g) Are there any nontrivial functions ϕ for which $\mathbf{h}\phi = \phi$?

Hint: Begin with the convolution theorem (7.19), and set $\psi = \phi$.

7. Show that $\mathbf{h}\left[(1-x^2)^{+1/2}\right] = -x$.
8. The results of the previous two problems can be generalized. Let $T_n(x)$ and $U_n(x)$ denote the Chebychev polynomials of the first and second kinds of degree n. For $n \geq 1$, show that

$$\mathbf{h}\left[\frac{T_n(x)}{\sqrt{1-x^2}}\right] = U_{n-1}(x) \quad \text{and that} \quad \mathbf{h}\left[\sqrt{1-x^2}\,U_{n-1}(x)\right] = -T_n(x).$$

9. (a) Show that $\mathbf{P}[\cos x] = \sin x$ and that $\mathbf{P}[\sin x] = -\cos x$.
(b) Determine $\mathbf{P}[\cos(nx)]$ and $\mathbf{P}[\sin(nx)]$ for all integers $n \geq 1$.
10. Solve the singular integral equation of the first kind with logarithmic kernel

$$\int_{-1}^{+1} \ln|x-t|\,\phi(t)\,\mathrm{d}t = \sqrt{1-x^2}.$$

Chapter 8
Linear Systems of Integral Equations

A *system of integral equations* is a set of two or more integral equations in two or more unknown functions. Usually, all of the equations belonging to a system are of the same type, but this need not be the case. Since linear systems of Fredholm, Volterra, or singular integral equations occur very commonly in practice, they are the subjects of this chapter.

There are a variety of methods for solving systems of integral equations:

- Converting a system of integral equations into a different type of system is an especially useful technique. For example, a linear system of integral equations can often be converted into a system of linear equations or a system of linear differential equations by differentiation or by employing a transform method. The solution to the converted system can then be related back to the solution of the system of integral equations.
- A linear system of integral equations can be converted into a single linear integral equation in several ways. One method of doing this involves writing the system as a single integral equation by employing matrix notation. The kernel is a single matrix function that incorporates all of the kernels in the system. The free term and the solution are vector functions. Then, an analogous version of the Theorem of Successive Substitution can be applied to the equation in matrix form.
- The integral equations in a linear system can be substituted into one another, yielding a single integral equation with a more complex kernel. The solutions to the more complex equation can then be related back to the solutions of the linear system.
- In some cases, a linear system of integral equations can be converted into a decoupled system of linear equations that can be solved individually.

In many systems of interest, these conversion methods are not applicable. Thus, the importance of determining approximate solutions to such systems of integral equations cannot be underestimated. It may be necessary at times to approximate one or more of the kernels or free terms in order to produce an approximate solution. Difficult questions of convergence and stability must be addressed in the ensuing analysis in order to verify accuracy.

S.M. Zemyan, *The Classical Theory of Integral Equations: A Concise Treatment*, 287
DOI 10.1007/978-0-8176-8349-8_8, © Springer Science+Business Media, LLC 2012

In Sect. 8.1, we present several tools of the trade which are indispensible for the comprehension of the material in this chapter.

In Sect. 8.2, we consider systems of Volterra integral equations. We show how to solve some simple systems when the kernels are of convolution type by employing the Laplace transform.

In Sect. 8.3, we investigate the relationship between linear systems of integral equations and differential equations of higher order.

In Sect. 8.4, we consider linear systems of Fredholm integral equations. We illustrate a method for solving them when the kernels are separable.

In Sect. 8.5, we consider systems of linear singular integral equations with Cauchy kernels or Hilbert kernels.

8.1 Tools of the Trade

In this chapter, the reader should be familiar with the following topics:

- *Norms of a continuous vector function*: If each $f_i(x)$ is continuous on the closed interval $[a,b]$, then we define the *norm of the continuous vector function* $\mathbf{f}(x) = (f_1(x),\ldots,f_n(x))$ by

$$\|\mathbf{f}(x)\| = \max_{1\le i\le n} |f_i(x)|\,. \tag{8.1}$$

  For each $x \in [a,b]$, vector norms satisfy the following axioms:

  1. $\|\mathbf{f}(x)\| \ge 0$
  2. $\|\mathbf{f}(x)\| = 0$ if and only if $\mathbf{f}(x) = \mathbf{0}$
  3. $\|\mathbf{f}(x) + \mathbf{g}(x)\| \le \|\mathbf{f}(x)\| + \|\mathbf{g}(x)\|$
  4. $\|k\mathbf{f}(x)\| = |k|\,\|\mathbf{f}(x)\|$ for any scalar k

  Since $\mathbf{f}(x)$ is continuous, there exists a constant $\mathcal{F}$ such that

$$\mathcal{F} = \max_{a\le x\le b} \|\mathbf{f}(x)\|.$$

  Also, it is readily shown that

$$\left\| \int_a^b \mathbf{f}(t)\,\mathrm{d}t \right\| \le \int_a^b \|\mathbf{f}(t)\|\,\mathrm{d}t\,. \tag{8.2}$$

- *The norm of a matrix*: Let $\mathbf{A}$ be either a real-valued or complex-valued matrix. The *norm of the matrix* $\mathbf{A}$, denoted by $\|\mathbf{A}\|$, is a nonnegative number with the following properties:

1. $\|\mathbf{A}\| \geq 0$.
2. $\|\mathbf{A}\| = 0$ if and only if $\mathbf{A}$ is the zero matrix.
3. $\|k\mathbf{A}\| = |k|\,\|\mathbf{A}\|$ for any scalar k.
4. $\|\mathbf{A}+\mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|$. (The triangle inequality holds.)
5. $\|\mathbf{AB}\| \leq \|\mathbf{A}\|\,\|\mathbf{B}\|$.

Matrix norms can be naturally induced from vector norms. For example, if $\|\mathbf{z}\|$ is the Euclidean vector norm, then the matrix norm naturally induced from it is defined by

$$\|\mathbf{A}\| = \max_{\|\mathbf{z}\|=1} \|\mathbf{Az}\|.$$

If $\mathbf{z}$ belongs to $\mathbb{R}^n$ and $\mathbf{A}$ is an $n \times n$ matrix, then $\mathbf{Az}$ belongs to $\mathbb{R}^n$. In this situation, the Euclidean norm of $\mathbf{A}$ is equal to the greatest length of all possible images of unit vectors.

- *Norms of a continuous matrix function*: A continuous matrix function of one or more variables is a matrix whose entries are continuous functions of one or more variables. The norm of a continuous matrix function replaces the concept of the absolute value in proofs of existence and uniqueness theorems for systems of integral equations written in matrix form.

  The norm of interest in this chapter is the *maximum absolute row sum norm* defined by

$$\|\mathbf{A}\| = \max_{1\leq i\leq n} \sum_{j=1}^{n} \left|a_{ij}\right|. \tag{8.3}$$

If each $k_{ij}(x,t)$ is continuous on the square $Q(a,b)$, then the norm of the matrix function $\mathbf{K}(x,t) = (k_{ij}(x,t))$ for each $(x,t) \in Q(a,b)$ is given by

$$\|\mathbf{K}(x,t)\| = \max_{1\leq i\leq n} \sum_{j=1}^{n} \left|k_{ij}(x,t)\right|.$$

Since each $\mathbf{k}_{ij}(x,t)$ is continuous, there exists a constant $\mathcal{K}$ such that

$$\mathcal{K} = \max_{Q(a,b)} \|\mathbf{K}(x,t)\|.$$

Also, it is readily shown for this norm that

$$\left\| \int_a^b \mathbf{K}(x,t)\,\mathrm{d}t \right\| \leq \int_a^b \|\mathbf{K}(x,t)\|\,\mathrm{d}t. \tag{8.4}$$

8.2 Systems of Volterra Equations

A system of linear Volterra integral equations can be rewritten in terms of matrix notation.

Consider the system consisting of three Volterra integral equations of the second kind in three unknowns:

$$\phi_1(x) = f_1(x) + \int_0^x k_{11}(x,t)\,\phi_1(x) + k_{12}(x,t)\,\phi_2(x) + k_{13}(x,t)\,\phi_3(x)\,\mathrm{d}t,$$

$$\phi_2(x) = f_2(x) + \int_0^x k_{21}(x,t)\,\phi_1(x) + k_{22}(x,t)\,\phi_2(x) + k_{23}(x,t)\,\phi_3(x)\,\mathrm{d}t,$$

$$\phi_3(x) = f_3(x) + \int_0^x k_{31}(x,t)\,\phi_1(x) + k_{32}(x,t)\,\phi_2(x) + k_{33}(x,t)\,\phi_3(x)\,\mathrm{d}t.$$

In matrix notation, this system appears in the alternate form

$$\begin{pmatrix} \phi_1(x) \\ \phi_2(x) \\ \phi_3(x) \end{pmatrix} = \begin{pmatrix} f_1(x) \\ f_2(x) \\ f_3(x) \end{pmatrix} + \int_0^x \begin{pmatrix} k_{11}(x,t) & k_{12}(x,t) & k_{13}(x,t) \\ k_{21}(x,t) & k_{22}(x,t) & k_{23}(x,t) \\ k_{31}(x,t) & k_{32}(x,t) & k_{33}(x,t) \end{pmatrix} \begin{pmatrix} \phi_1(t) \\ \phi_2(t) \\ \phi_3(t) \end{pmatrix} \mathrm{d}t,$$

or more compactly as

$$\mathbf{p}(x) = \mathbf{f}(x) + \int_0^x \mathbf{K}(x,t)\,\mathbf{p}(t)\,\mathrm{d}t,$$

where obvious assignments have been made.

It can be shown that this system has a unique solution $\mathbf{p}(x)$ on the interval $[0,b]$ if $\mathbf{f}(x)$ is continuous on the interval $[0,b]$ and if $\mathbf{K}(x,t)$ is continuous on the triangle $T(0,b)$. The proof of this fact is quite similar to the proof for a Volterra integral equation of the second kind in one variable, with matrix norms (see Sect. 8.1) replacing absolute values.

Some systems of Volterra integral equations can be quite difficult to solve, especially if the component kernels $k_{ij}(x,t)$ have very different functional forms. However, if each kernel is a convolution kernel, then the Laplace transform can be used very efficiently to solve the system.

Suppose that $k_{ij}(x,t) = m_{ij}(x-t)$ for all i and j. Set

$$F_i(s) = \mathcal{L}\{f_i(x)\}, \qquad M_{ij}(s) = \mathcal{L}\{m_{ij}(x)\}, \quad \text{and} \quad \Phi_i(s) = \mathcal{L}\{\phi_i(x)\}.$$

The Laplace transform converts the system of three integral equations into the following system of three linear equations, written as

$$\Phi_1(s) = F_1(s) + M_{11}(s)\,\Phi_1(s) + M_{12}(s)\,\Phi_2(s) + M_{13}(s)\,\Phi_3(s),$$

$$\Phi_2(s) = F_2(s) + M_{21}(s)\,\Phi_1(s) + M_{22}(s)\,\Phi_2(s) + M_{23}(s)\,\Phi_3(s),$$

$$\Phi_3(s) = F_3(s) + M_{31}(s)\,\Phi_1(s) + M_{32}(s)\,\Phi_2(s) + M_{33}(s)\,\Phi_3(s).$$

or more compactly in matrix form as

$$\mathbf{P}(s) = \mathbf{F}(s) + \mathbf{M}(s)\,\mathbf{P}(s),$$

where $\mathbf{P}(s) = (\Phi_1(s), \Phi_2(s), \Phi_3(s))^{\mathrm{T}}$, $\mathbf{F}(s) = (F_1(s), F_2(s), F_3(s))^{\mathrm{T}}$, and the matrix $\mathbf{M}(s) = (M_{ij}(s))$. It follows directly from this equation that

$$(\mathbf{I} - \mathbf{M}(s))\,\mathbf{P}(s) = \mathbf{F}(s),$$

from which we conclude that

$$\mathbf{P}(s) = (\mathbf{I} - \mathbf{M}(s))^{-1}\,\mathbf{F}(s),$$

provided that the inverse exists. The solution is then $\mathbf{p}(x) = \mathcal{L}^{-1}\{\mathbf{P}(s)\}$.

The introductory example above consists of a system of three equations in three unknown functions. Obviously, systems can consist of an arbitrary number of equations and unknown functions. Furthermore, the left-hand side of these equations can consist of a linear combination of the unknown functions as well. Thus, in general, a system of n integral equations of Volterra type can appear in the expanded form

$$\sum_{j=1}^{n} a_{ij}(x)\,\phi_j(x) = f_j(x) + \int_0^x \left(\sum_{j=1}^{n} k_{ij}(x,t)\,\phi_j(t) \right) \mathrm{d}t$$

for $i = 1, \ldots, n$. In the more compact matrix form, this equation becomes

$$\mathbf{A}(x)\,\mathbf{p}(x) = \mathbf{f}(x) + \int_0^x \mathbf{k}(x,t)\,\mathbf{p}(t)\,\mathrm{d}t,$$

where obvious assignments have been made. In applications, $\mathbf{A}(x)$ is assumed to be constant, bounded, nonsingular, or diagonalizable.

Illustrative Examples

- *Example 1*: Consider the linear system of Volterra integral equations

$$\phi(x) = 2 + \int_0^x (x-t)^2\, \psi(t)\, \mathrm{d}t,$$

$$\psi(x) = x + 16 \int_0^x (x-t)\, \phi(t)\, \mathrm{d}t.$$

The Laplace transform can be applied to this system of integral equations since the kernels are of convolution type.

If we set $\Phi(s) = \mathcal{L}\{\phi(x)\}$ and $\Psi(s) = \mathcal{L}\{\psi(x)\}$, then this system of integral equations is converted by the Laplace transform into the system of linear equations

$$\Phi(s) - \frac{1}{s^3}\Psi(s) = \frac{2}{s},$$
$$-\frac{16}{s^2}\Phi(s) + \Psi(s) = \frac{1}{s^2},$$

The solution of this linear system consists of the transforms

$$\Phi(s) = \frac{2s^4 + 2}{s^5 - 32}$$

and

$$\Psi(s) = \frac{s^3 + 32s^2}{s^5 - 32}.$$

Inversion is readily accomplished if these transforms are expanded into more recognizable forms.

If we set[1] $g = (\sqrt{5}+1)/2$, $h = 1/g = (\sqrt{5}-1)/2$, $a = \sqrt{(5+\sqrt{5})/2}$, and $b = \sqrt{(5-\sqrt{5})/2}$, then $\Phi(s)$ and $\Psi(s)$ can be decomposed as

$$\Phi(s) = \frac{17}{40}\left(\frac{1}{s-2}\right) + \frac{1}{80a}\left[\left(63+\sqrt{5}\right) a \left(\frac{s-h}{(s-h)^2+a^2}\right) - \left(5+\sqrt{5}\right)\left(\frac{a}{(s-h)^2+a^2}\right)\right]$$

[1] g is known as the Golden Ratio.

$$+\frac{1}{80b}\left[\left(63-\sqrt{5}\right)b\left(\frac{s+g}{(s+g)^2+b^2}\right)\right.$$
$$\left.-\left(5-\sqrt{5}\right)\left(\frac{b}{(s+g)^2+b^2}\right)\right]$$

and

$$\Psi(s)=\frac{17}{10}\left(\frac{1}{s-2}\right)$$
$$-\frac{1}{20a}\left[\left(17+15\sqrt{5}\right)a\left(\frac{s-h}{(s-h)^2+a^2}\right)-\left(5+33\sqrt{5}\right)\left(\frac{a}{(s-h)^2+a^2}\right)\right]$$
$$-\frac{1}{20b}\left[\left(17-15\sqrt{5}\right)b\left(\frac{s+g}{(s+g)^2+b^2}\right)-\left(5-33\sqrt{5}\right)\left(\frac{b}{(s+g)^2+b^2}\right)\right].$$

After inversion, we obtain the solutions

$$\phi(x)=\frac{17}{40}\mathrm{e}^{2x}+\frac{1}{80a}\mathrm{e}^{hx}\left[\left(63+\sqrt{5}\right)a\cos(ax)-(5+\sqrt{5})\sin(ax)\right]$$
$$+\frac{1}{80b}\mathrm{e}^{-gx}\left[\left(63-\sqrt{5}\right)b\cos(bx)-(5-\sqrt{5})\sin(bx)\right]$$

and

$$\psi(x)=\frac{17}{10}\mathrm{e}^{2x}-\frac{1}{20a}\mathrm{e}^{hx}\left[\left(17+15\sqrt{5}\right)a\cos(ax)-\left(5+33\sqrt{5}\right)\sin(ax)\right]$$
$$-\frac{1}{20b}\mathrm{e}^{-gx}\left[\left(17-15\sqrt{5}\right)b\cos(bx)-\left(5-33\sqrt{5}\right)\sin(bx)\right].$$

It is worth noting that this system of Volterra integral equations can be decoupled via differentiation into two initial value problems of the fifth order. Specifically, if we differentiate both integral equations five times and combine the results algebraically, then we find that $\phi(x)$ is the solution of the ordinary differential equation

$$\phi''''''(x)=32\,\phi(x)$$

together with the initial conditions

$$\phi(0)=\phi''''(0)=2\quad\text{and}\quad\phi'(0)=\phi''(0)=\phi'''(0)=0,$$

and that $\psi(x)$ is the solution of the ordinary differential equation

$$\psi''''''(x)=32\,\psi(x)$$

together with the initial conditions

$$\psi(0)=\psi'''(0)=\psi''''(0)=0,\qquad\psi'(0)=1,\quad\text{and}\quad\psi''(0)=32.$$

Although it is computationally possible to effect this conversion, it is rather cumbersome to actually carry it out. Hence, it is definitely not a recommended method in this case.

Section 8.2 Exercises

1. Solve the linear system of Volterra integral equations of the second kind given by

$$\phi(x) + 2\,\psi(x) = 2 + 2\int_0^x \phi(t)\,\mathrm{d}t,$$

$$\phi(x) - \psi(x) = 2x + \int_0^x (x-t)\,\psi(t)\,\mathrm{d}t.$$

Answer:

$$\phi(x) = \frac{4}{3}\mathrm{e}^{x} - \frac{2}{3}\mathrm{e}^{-x/6}\cos\left(\frac{\sqrt{23}}{6}x\right) + \frac{2\sqrt{23}}{23}\mathrm{e}^{-x/6}\sin\left(\frac{\sqrt{23}}{6}x\right)$$

and

$$\psi(x) = \frac{2}{3}\mathrm{e}^{x} - \frac{16\sqrt{23}}{69}\mathrm{e}^{-x/6}\sin\left(\frac{\sqrt{23}}{6}x\right).$$

Show that $\phi(x)$ satisfies the initial value problem that consists of the differential equation

$$3\,\phi'''(x) - 2\,\phi''(x) + \phi'(x) - 2\,\phi(x) = 0$$

together with the initial conditions

$$\phi(0) = \frac{2}{3}, \qquad \phi'(0) = \frac{16}{9}, \quad \text{and} \quad \phi''(0) = \frac{44}{27}.$$

Show also that $\psi(x)$ satisfies the initial value problem consisting of the differential equation

$$3\,\psi'''(x) - 2\,\psi''(x) + \psi'(x) - 2\,\psi(x) = 0$$

together with the initial conditions

$$\psi(0) = \frac{2}{3}, \qquad \psi'(0) = -\frac{2}{9}, \quad \text{and} \quad \psi''(0) = \frac{26}{27}.$$

In doing so, you have shown that the system of Volterra integral equations of the second kind can be successfully decoupled via differentiation.

2. Consider the matrix system of Volterra singular integral equations of the second kind

$$\mathbf{p}(x) = \mathbf{f}(x) + \int_0^x \mathbf{A}\,\mathbf{p}(t)\,\mathrm{d}t,$$

where

$$\mathbf{p}(x) = \begin{pmatrix} \phi_1(x) \\ \phi_2(x) \\ \phi_3(x) \end{pmatrix}, \qquad \mathbf{f}(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \\ f_3(x) \end{pmatrix}, \quad \text{and} \quad \mathbf{A} = \begin{pmatrix} 4 & 0 & 1 \\ -2 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}.$$

(a) If $\mathbf{f}(x) = \mathbf{0}$, then $\mathbf{p}(x) = \mathbf{0}$ is obviously a solution to this system. Is this solution unique?

(b) Solve the matrix system if

$$\mathbf{f}(x) = \mathbf{c}(x) = \begin{pmatrix} \cos x \\ \cos(2x) \\ \cos(3x) \end{pmatrix}$$

and if

$$\mathbf{f}(x) = \mathbf{s}(x) = \begin{pmatrix} \sin x \\ \sin(2x) \\ \sin(3x) \end{pmatrix}.$$

Hints: The matrix $\mathbf{A}$ is diagonalizable. The eigenvalues of $\mathbf{A}$ are $\lambda_1 = 1$, $\lambda_2 = 2$, and $\lambda_3 = 3$. The eigenvectors corresponding to these eigenvalues are

$$\mathbf{v}_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad \mathbf{v}_2 = \begin{pmatrix} -1 \\ 2 \\ 2 \end{pmatrix}, \quad \text{and} \quad \mathbf{v}_3 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix},$$

respectively.

Diagonalize $\mathbf{A}$ by setting $\mathbf{A} = \mathbf{PDP}^{-1}$, where $\mathbf{D}$ is the diagonal matrix with the eigenvalues of $\mathbf{A}$ as diagonal entries, and the columns of $\mathbf{P}$ are the eigenvectors of $\mathbf{A}$.

Decouple the system into three separate Volterra integral equations of the second kind by setting $\mathbf{q}(x) = \mathbf{P}^{-1}\mathbf{p}(x)$.

Partial answer: The solutions to both equations assume the form

$$\mathbf{p}(x) = c_1 \mathrm{e}^{x}\,\mathbf{v}_1 + c_2 \mathrm{e}^{2x}\,\mathbf{v}_2 + c_3 \mathrm{e}^{3x}\,\mathbf{v}_3 + \mathbf{C}\mathbf{c}(x) + \mathbf{S}\mathbf{s}(x),$$

where $\mathbf{C}$ and $\mathbf{S}$ are constant matrices.

Explain why this form could have been reasonably predicted before any attempt to solve the system.

3. Consider the matrix system of Volterra integral equations of the second kind

$$\mathbf{p}(x) = \mathbf{f}(x) + \int_0^x \mathbf{A}\,\mathbf{p}(t)\,\mathrm{d}t,$$

where

$$\mathbf{p}(x) = \begin{pmatrix} \phi(x) \\ \psi(x) \end{pmatrix}, \qquad \mathbf{f}(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \end{pmatrix}, \quad \text{and} \quad \mathbf{A} = \begin{pmatrix} 0 & 1 \\ 4 & 0 \end{pmatrix}.$$

(a) Solve this coupled system by converting it into a decoupled system of second-order differential equations.

(b) Solve this system again by converting it into a decoupled system of Volterra integral equations. Use the fact that $\mathbf{A}$ is diagonalizable. Its eigenvalues are $\lambda_1 = +2$ and $\lambda_2 = -2$, and its corresponding eigenvectors are $\mathbf{v}_1 = (1,2)^{\mathrm{T}}$ and $\mathbf{v}_2 = (1,-2)^{\mathrm{T}}$.
Answer: With $k = 1/(1-4\lambda^2)$, the solution is

$$\mathbf{p}(x) = k\left[2\mathrm{e}^{-x}\begin{pmatrix} -\lambda \\ 1 \end{pmatrix} + \mathrm{e}^{x}\begin{pmatrix} 1 \\ 4\lambda \end{pmatrix} - 4\lambda^2 \mathrm{e}^{2\lambda x}\,\mathbf{v}_1 + 2\lambda \mathrm{e}^{-2\lambda x}\,\mathbf{v}_2\right].$$

If $\lambda = \pm 1/2$, the solution exists as a limit.

8.3 Differential Equations

It is possible to convert a linear ordinary differential equation of higher order into a system of integral equations. Before we demonstrate this procedure, we recall some basic facts.

Valuable techniques for integrating higher derivatives were introduced in Sect. 5.1. These techniques were used in Sect. 5.2.1 to convert boundary value problems, consisting of a second-order ordinary differential equation together with two boundary conditions, into a single Fredholm integral equation. To be specific, it was shown that the second-order equation

$$\phi''(x) + \lambda\, \phi(x) = 0$$

together with the boundary conditions $\phi(0) = \phi(1) = 0$ can be converted to the homogeneous Fredholm integral equation

$$\phi(x) = \lambda \int_0^1 G(x,t)\, \phi(t)\, \mathrm{d}t$$

where the triangular kernel is prescribed by

$$G(x,t) = \begin{cases} \dfrac{t(1-x)}{1} & \text{if } 0 \le t \le x \le 1, \\ \dfrac{x(1-t)}{1} & \text{if } 0 \le x \le t \le 1. \end{cases}$$

Illustrative Examples

- *Example 1*: In this example, we convert the boundary value problem that consists of an ordinary differential equation of the fourth-order together with suitable boundary conditions into a system of integral equations.

  Consider the fourth order equation

$$\phi''''(x) - \lambda\, \phi(x) = 0 \tag{8.5}$$

  together with the boundary conditions

$$\phi(0) = \phi(1) = 0 \quad \text{and} \quad \phi''(0) = \phi''(1) = 0.$$

  This equation can be solved directly. If $\lambda_n = n^4\pi^4$, then the solution is $\phi_n(x) = c\, \sin(n\pi x)$, where c is an arbitrary constant. If $\lambda \neq n^4\pi^4$ for any integer n, then $\phi(x) \equiv 0$.

  Clearly, this fourth-order equation can be decomposed into two second-order equations by setting

$$\phi''(x) = -\psi(x) \quad \text{and} \quad \psi''(x) = -\lambda\, \phi(x)$$

  with the accompanying boundary conditions

$$\phi(0) = \phi(1) = 0 \quad \text{and} \quad \psi(0) = \psi(1) = 0.$$

  If we proceed as above, then these two equations can be converted into the system consisting of the *coupled* integral equations

$$\phi(x) = \int_0^1 G(x,t)\,\psi(t)\,\mathrm{d}t,$$

$$\psi(x) = \lambda \int_0^1 G(x,t)\,\phi(t)\,\mathrm{d}t.$$

After replacing x by t and t by s in the second equation and substituting it into the first equation, this system becomes

$$\phi(x) = \int_0^1 G(x,t) \left(\lambda \int_0^1 G(t,s)\,\phi(s)\,\mathrm{d}s \right) \mathrm{d}t.$$

After interchanging the order of integration, we obtain

$$\phi(x) = \lambda \int_0^1 \left(\int_0^1 G(x,t)\,G(t,s)\,\mathrm{d}t \right) \phi(s)\,\mathrm{d}s,$$

or what is the same, after replacing s by t,

$$\phi(x) = \lambda \int_0^1 G_2(x,t)\,\phi(t)\,\mathrm{d}t. \tag{8.6}$$

The iterated kernel $G_2(x,t)$ can be computed directly. If $0 \le x \le t \le 1$, then

$$G_2(x,t) = \left(\int_0^x + \int_x^t + \int_t^1 \right) G(x,s)\,G(s,t)\,\mathrm{d}s.$$

A similar expression holds if $0 \le t \le x \le 1$. Note that

$$\int_0^x G(x,s)\,G(s,t)\,\mathrm{d}s = \int_0^x s(1-x)s(1-t)\,\mathrm{d}s$$

$$= \frac{1}{3}x^3(1-x)(1-t),$$

that

$$\int_x^t G(x,s)\,G(s,t)\,\mathrm{d}s = \int_x^t x(1-s)s(1-t)\,\mathrm{d}s$$

$$= \frac{1}{6}x(1-t)\left(3t^2-3x^2-2t^3+2x^3\right),$$

and that

$$\int_t^1 G(x,s)\,G(s,t)\,\mathrm{d}s = \int_t^1 x(1-s)t(1-s)\,\mathrm{d}s$$

$$= \frac{1}{3}xt(1-t)^3.$$

It follows from these calculations that

$$G_2(x,t) = \begin{cases} \dfrac{1}{6}x(1-t)\left(2t-x^2-t^2\right) & \text{if } 0 \le x \le t \le 1, \\ \dfrac{1}{6}t(1-x)\left(2x-t^2-x^2\right) & \text{if } 0 \le t \le x \le 1. \end{cases}$$

We conclude from the analysis above that the solutions to the fourth-order differential equation (8.5) are among the eigenfunctions of $G_2(x,t)$, i.e., the solutions to the integral equation (8.6).

Since the kernel $G(x,t)$ and its iterate $G_2(x,t)$ are symmetric, it follows from Proposition 3.2.4 that the eigenvalues of $G_2(x,t)$ are the squares of the eigenvalues of $G(x,t)$. Furthermore, their eigenfunctions are the same.

In Illustrative Example 1 of Sect. 5.2.2, it was shown that the eigenvalues of the triangular kernel $G(x,t)$ are $n^2\pi^2$ and that its eigenfunctions are $\phi_n(x) = \sin(n\pi x)$. Hence, the eigenvalues of $G_2(x,t)$ are $n^4\pi^4$, and its eigenfunctions are the same.

Section 8.3 Exercises

1. Consider again the fourth-order differential equation

$$\phi''''(x) - \lambda\,\phi(x) = 0$$

that we discussed in Example 1. It was noted that this differential equation can be decomposed into the two second-order differential equations

$$\phi''(x) = -\psi(x) \quad \text{and} \quad \psi''(x) = -\lambda\,\phi(x).$$

Replace the boundary conditions given there with the similar conditions

$$\phi(0) = \phi'(0) = 0 \quad \text{and} \quad \psi(1) = \psi'(1) = 0.$$

Follow the methods introduced in Sect. 5.1, Sect. 5.2.1, and Example 1 in order to convert this pair of two ordinary differential equations into a system of two coupled integral equations. Substitute one of them into the other one, thereby obtaining a single integral equation with the symmetric kernel

$$K(x,t) = \begin{cases} \frac{1}{2}t^2x - \frac{1}{6}t^3 & \text{if } 0 \le t \le x \le 1, \\ \frac{1}{2}x^2t - \frac{1}{6}x^3 & \text{if } 0 \le x \le t \le 1. \end{cases}$$

Note that the kernel determined here is different from the kernel that was determined in Example 1. The point is that the form of the kernel depends upon the initial conditions that are specified for $\phi(x)$ and $\psi(x)$.

8.4 Systems of Fredholm Equations

A system of n linear integral equations of Fredholm type has the form

$$\sum_{j=1}^{n} a_{ij}(x)\,\phi_j(x) = f_j(x) + \int_a^b \left(\sum_{j=1}^{n} k_{ij}(x,t)\,\phi_j(t) \right) \mathrm{d}t$$

for $i = 1, \ldots, n$, where the functions $a_{ij}(x)$ and $f_j(x)$ are continuous on the interval $[a,b]$, and $k_{ij}(x,t)$ is continuous on $Q(a,b)$.

By employing the matrix notation described in Sect. 8.2, this system can be written in matrix form as

$$\mathbf{A}(x)\,\mathbf{p}(x) = \mathbf{f}(x) + \int_a^b \mathbf{K}(x,t)\,\mathbf{p}(t)\,\mathrm{d}t,$$

where obvious assignments have been made. If $\mathbf{A}(x) = \mathbf{0}$, then the system is said to be of the *first kind*. If $\mathbf{A}(x)$ is invertible, then it is said to be of the *second kind*. If $\mathbf{A}(x) \neq \mathbf{0}$ but it is not invertible, then it is said to be of the *third kind*. The matrix $\mathbf{A}(x)$ is often assumed to be constant, bounded, nonsingular, or diagonalizable.

Some linear systems of Fredholm integral equations can be quite difficult to solve, especially if the component kernels $k_{ij}(x,t)$ have different functional forms. However, if the kernels are separable, then the algebraic methods that were introduced in Chap. 1 can be used very efficiently to solve the system.

Illustrative Examples

- *Example 1*: Consider the linear system of Fredholm integral equations of the second kind:

$$\mathbf{A}(x)\,\mathbf{p}(x) = \mathbf{f}(x) + \int_a^b \mathbf{K}(x,t)\,\mathbf{p}(t)\,\mathrm{d}t,$$

where

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \qquad \mathbf{f}(x) = \begin{pmatrix} 40x \\ x \end{pmatrix}, \quad \mathbf{p}(x) = \begin{pmatrix} \phi(x) \\ \psi(x) \end{pmatrix},$$

and

$$\mathbf{K}(x,t) = \begin{pmatrix} -x-t & 0 \\ 0 & -xt \end{pmatrix}.$$

In expanded form, this system can be written as

$$\phi(x) + \psi(x) = 40x - \int_0^1 (x+t)\,\phi(t)\,\mathrm{d}t,$$

$$\phi(x) - \psi(x) = x - \int_0^1 xt\,\psi(t)\,\mathrm{d}t.$$

Note that the kernel elements $k_{11}(x,t)$ and $k_{22}(x,t)$ are continuous on the square $Q(0,1)$ and separable. The element $k_{11}(x,t)$ is the sum of two terms, and the element $k_{22}(x,t)$ is a single term. Thus, we should expect the representative forms of the solutions to involve three constants which are to be determined.

If we set

$$c_1 = \int_0^1 \phi(t)\,\mathrm{d}t, \qquad c_2 = \int_0^1 t\,\phi(t)\,\mathrm{d}t, \quad \text{and} \quad d = \int_0^1 t\,\psi(t)\,\mathrm{d}t,$$

then the system can be written in the alternate form

$$\phi(x) + \psi(x) = (40 - c_1)\,x - c_2,$$

$$\phi(x) - \psi(x) = (1-d)\,x,$$

from which we conclude that $\phi(x)$ and $\psi(x)$ must assume the necessarily linear forms

$$\phi(x) = \frac{1}{2}\,(41 - c_1 - d)\,x - \frac{1}{2}\,c_2$$

and

$$\psi(x) = \frac{1}{2}(39 - c_1 + d)\,x - \frac{1}{2}c_2.$$

If we substitute $\phi(x)$ into the definitions of c_1 and c_2, and $\psi(x)$ into the definition of d, then we obtain the linear system

$$\begin{aligned} 5c_1 + 2c_2 + d &= 41, \\ 2c_1 + 15c_2 + 2d &= 82, \\ 2c_1 + 3c_2 + 10d &= 78. \end{aligned}$$

The unique solution to this linear system is given by

$$c_1 = \frac{11}{2}, \qquad c_2 = 4, \quad \text{and} \quad d = \frac{11}{2}.$$

Thus,

$$\phi(x) = 15x - 2$$

and

$$\psi(x) = \frac{39}{2}x - 2.$$

In matrix form, we can write the solution as

$$\mathbf{p}(x) = \begin{pmatrix} \phi(x) \\ \psi(x) \end{pmatrix} = \begin{pmatrix} 15x - 2 \\ \frac{39}{2}x - 2 \end{pmatrix}.$$

- *Example 2*: In this example, we will examine the Method of Successive Substitution for matrix systems of linear Fredholm integral equations of the second kind.

  If the matrix system

$$\mathbf{p}(x) = \mathbf{f}(x) + \lambda \int_0^1 \mathbf{K}(x,t)\,\mathbf{p}(t)\,\mathrm{d}t$$

is substituted into itself, then we obtain

$$\mathbf{p}(x) = \mathbf{f}(x) + \lambda \int_0^1 \mathbf{K}(x,t)\,\mathbf{f}(t)\,\mathrm{d}t + \lambda^2 \int_0^1 \mathbf{K}_2(x,t)\,\mathbf{p}(t)\,\mathrm{d}t,$$

where we have set

$$\mathbf{K}_2(x,t) = \int_0^1 \mathbf{K}(x,s)\,\mathbf{K}(s,t)\,\mathrm{d}t.$$

Additional iterations lead to the general form

$$\mathbf{p}(x) = \mathbf{f}(x) + \lambda \int_0^1 \left(\sum_{m=1}^{n} \lambda^{m-1} \mathbf{K}_m(x,t) \right) \mathbf{f}(t)\,\mathrm{d}t$$

$$+ \lambda^{n+1} \int_0^1 \mathbf{K}_{n+1}(x,t)\,\mathbf{p}(t)\,\mathrm{d}t,$$

where

$$\mathbf{K}_m(x,t) = \int_0^1 \mathbf{K}_{m-1}(x,s)\,\mathbf{K}(s,t)\,\mathrm{d}t.$$

Since $\mathbf{K}(x,s)$ need not commute with $\mathbf{K}(s,t)$ or its iterations, the order of matrix multiplication is important here.

The goal of our investigation is to show that there is a unique solution to the matrix system of the form

$$\mathbf{p}(x) = \mathbf{f}(x) + \lambda \int_0^1 \mathbf{R}(x,t;\lambda)\,\mathbf{f}(t)\,\mathrm{d}t,$$

where the resolvent matrix is given by the convergent series

$$\mathbf{R}(x,t;\lambda) = \sum_{m=1}^{\infty} \lambda^{m-1} \mathbf{K}_m(x,t).$$

Part of the proof consists in showing that the remainder tends to 0, i.e., that

$$\max_{0 \le x \le 1} \left\| \lambda^{n+1} \int_0^1 \mathbf{K}_{n+1}(x,t)\,p(t)\,\mathrm{d}t \right\| \to 0,$$

or, what is the same, that

$$\max_{a \le x \le b} \left\| \mathbf{p}(x) - \mathbf{f}(x) - \lambda \int_0^1 \left(\sum_{m=1}^{n} \lambda^{m-1} \mathbf{K}_m(x,t) \right) \mathbf{f}(t)\,\mathrm{d}t \right\| \to 0$$

as $n \to +\infty$.

The norm used above is the vector norm defined in Sect. 8.1 by Eq. (8.1), since the integrand is a continuous vector function of x.

The matrix norm used below for the matrix $\mathbf{K}(x,t)$ and its iterates is the maximum absolute row sum norm defined by Eq. (8.3).

By inequality (8.2) and the properties of matrix norms, we have

$$\left\| \lambda^{n+1} \int_0^1 \mathbf{K}_{n+1}(x,t)\, p(t)\, \mathrm{d}t \right\| \le \left|\lambda^{n+1}\right| \int_0^1 \left\|\mathbf{K}_{n+1}(x,t)\, \mathbf{p}(t)\right\| \mathrm{d}t$$

$$\le \left|\lambda^{n+1}\right| \int_0^1 \left\|\mathbf{K}_{n+1}(x,t)\right\| \left\|\mathbf{p}(t)\right\| \mathrm{d}t.$$

Since $\mathbf{K}(x,t)$ is a continuous matrix function, there exists a constant $\mathcal{K}$ such that

$$\mathcal{K} = \max_{Q(a,b)} \|\mathbf{K}(x,t)\|.$$

By inequality (8.4) and the properties of matrix norms, we have

$$\|\mathbf{K}_2(x,t)\| = \left\| \int_0^1 \mathbf{K}(x,s)\, \mathbf{K}(s,t)\, \mathrm{d}s \right\|$$

$$\le \int_0^1 \|\mathbf{K}(x,s)\, \mathbf{K}(s,t)\|\, \mathrm{d}s$$

$$\le \int_0^1 \|\mathbf{K}(x,s)\|\, \|\mathbf{K}(s,t)\|\, \mathrm{d}s$$

$$\le \mathcal{K}^2.$$

By a simple induction argument, it follows that $\|\mathbf{K}_m(x,t)\| \le \mathcal{K}^m$ for all $m \ge 1$ and all $(x,t) \in Q(0,1)$. Also, since $\mathbf{p}(x)$ is a continuous vector function, there exists a constant $\mathcal{P}$ such that

$$\mathcal{P} = \max_{a \le x \le b} \|\mathbf{p}(x)\|.$$

By combining the inequalities above, it now follows that

$$\max_{0 \le x \le 1} \left\| \lambda^{n+1} \int_0^1 \mathbf{K}_{n+1}(x,t)\, p(t)\, \mathrm{d}t \right\| \le (|\lambda|\,\mathcal{K})^{n+1}\, \mathcal{P} \to 0$$

as $n \to +\infty$, whenever $|\lambda| < 1/\mathcal{K}$. The student is asked to continue the discussion in Exercise 4.

Section 8.4 Exercises

1. Consider the matrix system of Fredholm integral equations of the second kind

$$\mathbf{p}(x) = \mathbf{f}(x) + \int_0^{\pi/2} \mathbf{A}\,\mathbf{p}(t)\,\mathrm{d}t,$$

where

$$\mathbf{p}(x) = \begin{pmatrix} \phi_1(x) \\ \phi_2(x) \\ \phi_3(x) \end{pmatrix}, \qquad \mathbf{f}(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \\ f_3(x) \end{pmatrix}, \quad \text{and} \quad \mathbf{A} = \begin{pmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{pmatrix}.$$

   (a) Explain why the solution has the form $\mathbf{p}(x) = \mathbf{f}(x) + \mathbf{c}$, where $\mathbf{c}$ is a constant vector.
   (b) Determine $\mathbf{c}$, thereby solving the system.
   (c) For general $\mathbf{f}$ and $\mathbf{A}$, does this system always have a solution?

2. Solve the system of Fredholm integral equations of the second kind

$$\phi(x) = 2 + \int_0^1 (x-t)^2\,\psi(t)\,\mathrm{d}t,$$

$$\psi(x) = x + 16\int_0^1 (x-t)\,\phi(t)\,\mathrm{d}t.$$

   Hint: By inspection, the solutions must assume the forms

$$\begin{aligned} \phi(x) &= d_0 x^2 - 2d_1 x + (2 + d_2), \\ \psi(x) &= (1 + 16c_0)\,x - 16c_1, \end{aligned}$$

   so that there are five undetermined constants in the solutions.
   Answer:

$$\begin{aligned} \phi(x) &= \frac{1}{44}\left(846x^2 - 1464x + 679\right), \\ \psi(x) &= \frac{9}{11}\left(103x - 28\right). \end{aligned}$$

   Note: This system is similar to the system of Volterra integral equations of the second kind that was considered in Example 1 in Sect. 8.2. The only difference between these systems is that the upper limit x in the Volterra system is replaced by 1 in the Fredholm system. However, the solutions to these systems are quite different.

3. In Example 1 above, three unknown constants were to be determined, and in the previous exercise, there were five. In general, how many unknown constants need to be determined in order to solve systems of this type?

   Hint: Suppose that the linear system of Fredholm integral equations of the second kind under consideration consists of N equations in N unknown functions $\phi_n(x)$. In this case, the corresponding matrix kernel $\mathbf{K}(x,t) = (k_{ij}(x,t))$ is of size $N \times N$. You may assume that every element $k_{ij}(x,t)$ of the kernel matrix is continuous on the square $Q(0,1)$ and is separable, with the usual form

$$k_{ij}(x,t) = \sum_{k=1}^{N_{ij}} a_{ijk}(x)\, b_{ijk}(t).$$

4. If $\mathbf{A}(t)$ is a continuous matrix function, show that

$$\left\| \int_0^1 \mathbf{A}(t)\,\mathrm{d}t \right\| \leq \int_0^1 \|\mathbf{A}(t)\|\,\mathrm{d}t,$$

   where the matrix here is the maximum absolute row sum norm.
5. State and prove a Theorem of Successive Substitution for matrix systems of Fredholm integral equations, analogous to Theorem 2.2.1 for a single Fredholm integral equation. You will need to show that the sequence defined by

$$\mathbf{p}_n(x) = \mathbf{f}(x) + \lambda \int_0^1 \left(\sum_{m=1}^{n} \lambda^{m-1}\, \mathbf{K}_m(x,t) \right) \mathbf{f}(t)\,\mathrm{d}t$$

   is a Cauchy sequence, and to explain why the solution is unique.
6. Solve the matrix system of Fredholm integral equations of the second kind

$$\mathbf{p}(x) = \mathbf{f}(x) + \int_0^{\pi} \mathbf{A}\,\mathbf{p}(t)\,\mathrm{d}t,$$

   where

$$\mathbf{p}(x) = \begin{pmatrix} \phi_1(x) \\ \phi_2(x) \\ \phi_3(x) \end{pmatrix}, \qquad \mathbf{f}(x) = \begin{pmatrix} 1 \\ \sin x \\ \cos x \end{pmatrix}, \quad \text{and} \quad \mathbf{A} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

   There is no need to diagonalize $\mathbf{A}$ here. The solution has the form $\mathbf{p}(x) = \mathbf{f}(x) + \mathbf{c}$, where $\mathbf{c}$ is a constant vector.

8.5 Systems of Singular Equations

A linear system of n singular integral equations in n unknown functions $\phi_j(x)$ with a Cauchy kernel has the general form

$$\sum_{j=1}^{n} a_{ij}(x)\,\phi_j(x) = f_i(x) + \int_{-1}^{1} \frac{1}{x-t}\left(\sum_{j=1}^{n} b_{ij}(x,t)\,\phi_j(t)\right) \mathrm{d}t + \int_{-1}^{1} \left(\sum_{j=1}^{n} k_{ij}(x,t)\,\phi_j(t)\right) \mathrm{d}t$$

for $i = 1,\ldots,n$, where the functions $a_{ij}(x)$ and $f_j(x)$ are assumed to satisfy a Lipschitz condition on the interval $[-1,1]$, and $k_{ij}(x,t)$ is assumed to satisfy a Lipschtiz condition in each of its variables.

This system[2] can be written in matrix form as

$$\mathbf{A}(x)\,\mathbf{p}(x) = \mathbf{f}(x) + \int_{-1}^{1} \frac{1}{x-t}\,\mathbf{B}(x,t)\,\mathbf{p}(t)\,\mathrm{d}t + \int_{-1}^{1} \mathbf{K}(x,t)\,\mathbf{p}(t)\,\mathrm{d}t,$$

where obvious assignments have been made. The matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{K}$ are of size $n \times n$. If $\mathbf{A}(x) = \mathbf{0}$, then the system is said to be of the *first kind*. If $\mathbf{A}(x)$ is invertible, it is said to be of the *second kind*, and if $\mathbf{A}(x) \neq \mathbf{0}$ but it is not invertible, then it is said to be of the *third kind*.

Some comments about the structure of this system are appropriate:

- A linear system of singular integral equations can be converted to a linear system of Fredholm equations, but the process can be troublesome.
- In practice, n is usually 2 or 3 since systems of this size arise when modeling physical systems in two or three dimensions.
- The system is defined on the open interval $(-1,+1)$ instead of the closed interval $[-1,+1]$ because the solutions $\phi_j(x)$ often have singularities at the endpoints ± 1 of the interval. Solutions with specifically prescribed behaviors near the singularities at the endpoints are often sought.
- The matrix $\mathbf{B}(x,t)$ is often assumed to be constant, bounded, nonsingular, or diagonalizable.
- The kernel $\mathbf{K}(x,t)$ is usually assumed to be bounded, but in general its elements may have singularities, often of logarithmic type. In this case, the terms containing singularities are separated from the others.

[2]For a general discussion of these systems, see Erdogan, F., *Approximate Solutions of Systems of Singular Integral Equations*, SIAM J. Appl. Math., Vol. 17, No. 6, November 1969.

A linear system of singular equations with the Hilbert kernel has the similar form

$$\sum_{j=1}^{n} a_{ij}(x)\,\phi_j(x) = f_i(x) + \int_{-1}^{1} \cot\left(\frac{t-x}{2}\right) \left(\sum_{j=1}^{n} b_{ij}(x,t)\,\phi_j(t)\right) \mathrm{d}t + \int_{-1}^{1} \left(\sum_{j=1}^{n} k_{ij}(x,t)\,\phi_j(t)\right) \mathrm{d}t$$

for $i = 1,\ldots,n$, where the functions $a_{ij}(x)$ and $f_j(x)$ are assumed to satisfy a Lipschitz condition on the interval $[-1,1]$, and $k_{ij}(x,t)$ is assumed to satisfy a Lipschtiz condition in each of its variables.

By employing the matrix notation described in Sect. 8.2, this system can be written in matrix form as

$$\mathbf{A}(x)\,\mathbf{p}(x) = \mathbf{f}(x) + \int_{-1}^{1} \cot\left(\frac{t-x}{2}\right) \mathbf{B}(x,t)\,\mathbf{p}(t)\,\mathrm{d}t + \int_{-1}^{1} \mathbf{K}(x,t)\,\mathbf{p}(t)\,\mathrm{d}t,$$

where obvious assignments have been made. The matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{K}$ are of size $n \times n$. If $\mathbf{A}(x) = \mathbf{0}$, then the system is said to be of the *first kind*. If $\mathbf{A}(x)$ is invertible, it is said to be of the *second kind*, and if $\mathbf{A}(x) \neq \mathbf{0}$ but it is not invertible, then it is said to be of the *third kind*.

Illustrative Examples

- *Example 1*: Some linear systems of singular integral equations can be decoupled.

  Consider the linear system of singular integral equations of the second kind in the matrix form

  $$\mathbf{A}\,\mathbf{p}(x) = \mathbf{f}(x) + \int_{-1}^{+1} \frac{1}{t-x}\,\mathbf{B}\,\mathbf{p}(t)\,\mathrm{d}t,$$

  where $\mathbf{A}$ and $\mathbf{B}$ are constant $n \times n$ matrices, $\mathbf{p}(x) = (\phi_1(x),\ldots,\phi_n(x))^{\mathrm{T}}$, and $\mathbf{f}(x) = (f_1(x),\ldots,f_n(x))^{\mathrm{T}}$.

  Suppose that $\mathbf{A}$ is nonsingular and $\mathbf{C} = \mathbf{A}^{-1}\mathbf{B}$ is diagonalizable. If we left-multiply the equation by $\mathbf{A}^{-1}$, then we obtain the equivalent system

  $$\mathbf{p}(x) = \mathbf{A}^{-1}\,\mathbf{f}(x) + \int_{-1}^{+1} \frac{1}{t-x}\,\mathbf{C}\,\mathbf{p}(t)\,\mathrm{d}t.$$

Since $\mathbf{C}$ is assumed to be diagonalizable, there exist matrices $\mathbf{D}$ and $\mathbf{P}$ such that $\mathbf{C} = \mathbf{PDP}^{-1}$, where $\mathbf{D}$ is a diagonal matrix whose diagonal entries are the n eigenvalues λ_i of $\mathbf{C}$ and $\mathbf{P} = (p_{ij})$ is an invertible matrix whose columns $\mathbf{e}_i$ are the n eigenvectors corresponding to these eigenvalues. With this in mind, the system can be rewritten as

$$\mathbf{p}(x) = \mathbf{A}^{-1}\mathbf{f}(x) + \int_{-1}^{+1} \frac{1}{t-x}\,\mathbf{PDP}^{-1}\mathbf{p}(t)\,\mathrm{d}t.$$

Now left-multiply this equation by $\mathbf{P}^{-1}$. If we define

$$\mathbf{q}(x) = \mathbf{P}^{-1}\mathbf{p}(x) = (\psi_1(x), \ldots, \psi_n(x))^{\mathrm{T}}$$

and

$$\mathbf{g}(x) = \mathbf{P}^{-1}\mathbf{A}^{-1}\mathbf{f}(x) = (g_1(x), \ldots, g_n(x))^{\mathrm{T}},$$

then the system becomes

$$\mathbf{q}(x) = \mathbf{g}(x) + \int_{-1}^{+1} \frac{1}{t-x}\,\mathbf{D}\,\mathbf{q}(t)\,\mathrm{d}t.$$

Since $\mathbf{D}$ is diagonal, we have successfully decoupled the system into a set of n individual singular equations of the second kind that can be explicitly written in the form

$$\psi_i(x) = g_i(x) + \lambda_i \int_{-1}^{+1} \frac{1}{t-x}\,\psi_i(t)\,\mathrm{d}t$$

for $i = 1, \ldots, n$.

After these n equations are solved for their individual solutions $\psi_i(x)$, the solution to the original system can be constructed, since $\mathbf{p}(x) = \mathbf{Pq}(x)$. Explicitly, we can write each $\phi_i(x)$ as a linear combination of the form

$$\phi_i(x) = \sum_{j=1}^{n} p_{ij}\,\psi_j(x).$$

For example, suppose that

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 7 \\ 0 & 1 & 2 \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} 0 & 1 & 1 \\ 11 & 7 & 4 \\ 3 & 2 & 1 \end{pmatrix}.$$

Then

$$\mathbf{C} = \mathbf{A}^{-1}\mathbf{B} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

The eigenvalues of $\mathbf{C}$ are $\lambda_1 = 2$ and $\lambda_2 = \lambda_3 = -1$. Corresponding directly to these eigenvalues, the eigenvectors of $\mathbf{C}$ are $\mathbf{e}_1 = (1,1,1)^{\mathrm{T}}$, $\mathbf{e}_2 = (1,0,-1)^{\mathrm{T}}$, and $\mathbf{e}_3 = (0,1,-1)^{\mathrm{T}}$. The matrix $\mathbf{P}$ is constructed with these eigenvectors as its columns, so that

$$\mathbf{P} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & -1 \end{pmatrix}.$$

Also,

$$\mathbf{g}(x) = \mathbf{P}^{-1}\mathbf{A}^{-1}\mathbf{f}(x) = \frac{1}{3}\begin{pmatrix} 1 & 1 & -3 \\ 2 & -1 & 3 \\ -1 & 5 & -18 \end{pmatrix}\mathbf{f}.$$

The decoupled system can now be explicitly written as the three separate equations

$$\psi_1(x) = \frac{1}{3}\left(f_1(x) + f_1(x) - 3f_3(x)\right) + 2\int_{-1}^{+1} \frac{1}{t-x}\,\psi_1(t)\,\mathrm{d}t,$$

$$\psi_2(x) = \frac{1}{3}\left(2f_1(x) - f_1(x) + 3f_3(x)\right) - \int_{-1}^{+1} \frac{1}{t-x}\,\psi_2(t)\,\mathrm{d}t,$$

$$\psi_3(x) = \frac{1}{3}\left(-f_1(x) + 5f_1(x) - 18f_3(x)\right) - \int_{-1}^{+1} \frac{1}{t-x}\,\psi_3(t)\,\mathrm{d}t.$$

Once $\mathbf{q}(x) = \mathbf{P}^{-1}\mathbf{p}(x) = \left(\psi_1(x),\ldots,\psi_n(x)\right)^{\mathrm{T}}$ has been determined, the solution $\mathbf{p}(x) = \left(\phi_1(x),\ldots,\phi_n(x)\right)^{\mathrm{T}}$ can be determined from it, since $\mathbf{p}(x) = \mathbf{P}\mathbf{q}(x)$. Explicitly, we may write the solution as

$$\begin{aligned} \phi_1(x) &= \psi_1(x) + \psi_2(x), \\ \phi_2(x) &= \psi_1(x) + \psi_3(x), \\ \phi_3(x) &= \psi_1(x) - \psi_2(x) - \psi_3(x). \end{aligned}$$

Appendix A
The 2010 Mathematics Subject Classification for Integral Equations

The classification that appears below (45–XX Integral Equations) is a printed form of part of MSC2010, a revision of the Mathematics Subject Classification (MSC) produced jointly by the editorial staffs of Mathematical Reviews (MR), a division of the American Mathematical Society (AMS), and of ZBMATH, formerly Zentralblatt für Mathematik (Zbl), in consultation with the mathematical community. It can be accessed in its entirety online at http://msc2010.org. It is reproduced here with the expressed permission of the American Mathematical Society.

45–XX	**Integral Equations**
45–00	General reference works (handbooks, dictionaries, bibliographies, etc.)
45–01	Instructional exposition (textbooks, tutorial papers, etc.)
45–02	Research exposition (monographs, survey articles)
45–03	Historical (must also be assigned at least one classification number from Section 01)
45–04	Explicit machine computation and programs (not the theory of computation or programming)
45–06	Proceedings, conferences, collections, etc.
45Axx	**Linear integral equations**
45A05	Linear integral equations
45A99	None of the above, but in this section
45Bxx	**Fredholm integral equations**
45B05	Fredholm integral equations
45B99	None of the above, but in this section
45Cxx	**Eigenvalue problems [See also 34Lxx, 35Pxx, 45P05, 47A75]**
45C05	Eigenvalue problems [See also 34Lxx, 35Pxx, 45P05, 47A75]
45C99	None of the above, but in this section

S.M. Zemyan, *The Classical Theory of Integral Equations: A Concise Treatment*,
DOI 10.1007/978-0-8176-8349-8, © Springer Science+Business Media, LLC 2012

45Dxx Volterra integral equations [See also 34A12]
45D05 Volterra integral equations [See also 34A12]
45D99 None of the above, but in this section
45Exx Singular integral equations [See also 30E20, 30E25, 44A15, 44A35]
45E05 Integral equations with kernels of Cauchy type [See also 35J15]
45E10 Integral equations of the convolution type (Abel, Picard, Toeplitz and Wiener-Hopf type) [See also 47B35]
45E99 None of the above, but in this section
45Fxx Systems of linear integral equations
45F05 Systems of nonsingular linear integral equations
45F10 Dual, triple, etc., integral and series equations
45F15 Systems of singular linear integral equations
45F99 None of the above, but in this section
45Gxx Nonlinear integral equations [See also 47H30, 47Jxx]
45G05 Singular nonlinear integral equations
45G10 Other nonlinear integral equations
45G15 Systems of nonlinear integral equations
45G99 None of the above, but in this section
45Hxx Miscellaneous special kernels [See also 44A15]
45H05 Miscellaneous special kernels [See also 44A15]
45H99 None of the above, but in this section
45Jxx Integro-ordinary differential equations [See also 34K05, 34K30, 47G20]
45J05 Integro-ordinary differential equations [See also 34K05, 34K30, 47G20]
45J99 None of the above, but in this section
45Kxx Integro-partial differential equations [See also 34K30, 35R09, 35R10, 47G20]
45K05 Integro-partial differential equations [See also 34K30, 35R09, 34R10, 47G20]
45K99 None of the above, but in this section
45Lxx Theoretical approximation of solutions [For numerical analysis, see 65Rxx]
45L05 Theoretical approximation of solutions [For numerical analysis, see 65Rxx]
45L99 None of the above, but in this section

45Mxx Qualitative behavior
45M05 Asymptotics
45M10 Stability theory
45M15 Periodic solutions
45M20 Positive solutions
45M99 None of the above, but in this section
45Nxx Abstract integral equations, integral equations in abstract spaces
45N05 Abstract integral equations, integral equations in abstract spaces
45N99 None of the above, but in this section
45Pxx Integral operators [See also 47B38, 47G10]
45P05 Integral operators [See also 47B38, 47G10]
45P99 None of the above, but in this section
45Qxx Inverse problems
45Q05 Inverse problems
45Q99 None of the above, but in this section
45Rxx Random integral equations [See also 60H20]
45R05 Random integral equations [See also 60H20]
45R99 None of the above, but in this section

Appendix B
Specialized Vocabularies and Sample Translations

A great deal of the classical theory of integral equations was originally published in French, Italian, and German, as well as in English.

In order to assist the student who may wish to consult the original sources for the theory, specialized vocabularies consisting of approximately 100 important terms are listed below. To illustrate the usage of these vocabularies, sample translations of the first two pages of each of three articles written in French, Italian, and German, respectively, are reproduced below. The citations for the full text of all three of these articles, each of which is available in its entirety on the SpringerLink web site, are:

Ghermanescu, M. Michel, Equations intégrales aux deux limites variables, *Annali di Mathematica Pura ed Applicata*, Volume 54, Number 1, 33–36

Tricomi, Francesco G., Equazioni integrali singolari del tipo di Carleman, *Annali di Mathematica Pura ed Applicata (4)*, Volume **39** (1955), 229–244.

Dinghas, A., Zur Existenz von Fixpunkten bei Abbildungen vom Abel-Liouvilleschen Typus, *Math. Zeitschr.*, Bd. 70, S. 174–189 (1958).

A portion of each of these three articles is reprinted here with kind permission from Springer Science + Business Media B.V. All three of the translations of the selections that appear below were done by Stephen M. Zemyan.

S.M. Zemyan, *The Classical Theory of Integral Equations: A Concise Treatment*,
DOI 10.1007/978-0-8176-8349-8, © Springer Science+Business Media, LLC 2012

B.1 French

B.1.1 Specialized French Vocabulary

English	French
adjoint	adjointe
alternative	alternative (*f*)
analytic	analytique
approximation theorem	théorème (*m*) d'approximation
bounded	borné
bilinear form	forme (*m*) bilinéare
boundary value problem	problème (*m*) marginal
Cauchy sequence	suite (*f*) de Cauchy
Cauchy principal value	valeur (*f*) principale de Cauchy
coefficient	coefficient (*m*)
compact	compact
complete	complet
complete orthonormal system	système (*m*) orthonormal complet
conjugate	conjugué
continuity	continuité (*f*)
continuous	continu
contraction	contraction (*f*)
contraction map	application (*f*) restringent
convergence	convergence (*f*)
convergent in the mean	convergent en moyenne
cosine	cosinus
cubic	cubique
determinant	déterminant (*f*)
differentiable	dérivable
differential equation	équation (*f*) différentielle
distance	distance (*f*)
eigenfunction	fonction (*f*) propre
eigenvalue	valeur (*f*) propre
eigenvector	vecteur (*m*) propre
finite	fini
first	premier, -ière
fixed point theorem	théorème (*m*) du point fixe
formula	formule (*f*)
Fourier transformation	transformation (*f*) de Fourier

English	French
fourth	quatrième
function	fonction (*f*)
gamma function	fonction (*f*) gamma
Hadamard's inequality	théorème (*m*) du déterminant de Hadamard
Hermitian form	forme (*f*) hermitienne
incomplete	incomplète
independent	indèpendant
inequality	inégalité (*f*)
infinite	infini
infinite series	série (*f*), progression (*f*)
initial value problem	problème (*m*) des conditions initiales
integrable	intègrable
integral equation	équation (*f*) intégrale
integration	intégration (*f*)
integration by parts	intégration (*f*) par parties
iterated	itére
kernel	noyau (*m*)
linear	linéaire
linearly dependent	linéairement independant
Lipschitz condition	condition (*f*) de Lipschitz
logarithmic	logarithmique
matrix	matrice (*f*)
method	méthode (*f*)
metric	métrique (*f*)
metric space	espace (*f*) métrique
necessary	nécessaire
norm	norme (*f*)
normal	normale
orthogonal	orthogonal
orthogonalization	procédé (*m*) d'orthogonalisation
orthonormal system	système (*m*) orthonormale
point	point (*m*)
polynomial	polynôme (*m*)
positive	positif

English	French
positive definite	positif defini
product kernel	noyau (*m*) produit
proof	démonstration (*f*)
rectangle	rectangle (*m*)
recursive	récursif
resolvent	résolvente (*f*)
Riemann integral	intégrale (*f*) de Riemann
second	deuxième
series	série (*f*)
separable	séparable
sequence	suite, séquence (*f*)
singular	singulier
skew	antisymmétrique
solution	solution (*f*)
spectrum	spectre (*m*)
square	carré (*m*)
square integrable	de carré integrable
substitution	substitution (*f*)
successive	successive
sufficient	suffisant
sum	somme (*f*)
symmetric	symétrique
term	terme (*m*)
third	troisième
transformation	transformation (*f*)
transpose matrix	matrice (*f*) transposée
trapezoid rule	méthode (*f*) des trapèze
trigonometric	trigonométrique
vector	vecteur (*m*)
zero	zéro (*m*)
zero of a polynomial	zéro (*m*) d'un polynôme

B.1.2 Sample Translation from French into English

Annali di Mathematica Pura ed Applicata
Volume 54, Number 1, 33 – 36.

Equations intégrales aux deux limites variables

par M. MICHEL GHERMANESCU (à Bucharest, Romania)

À *M. Enrico Bompiani pour son Jubilé scientifique*

1. Les équations intégrales à limites variables, découvertes par VITO VOLTERRA à la fin du siècle dernier, constituent l'un des puissants instruments de recherche mathématique moderne, grâce auquel le développement de la Physique mathématique et ceux de quelques autres domaines de recherche ont pris un essor considérable.

   Parmi ces équations, celles avec les deux limites variables, comprise dans le type général linéaire

$$p\,\varphi(x) = \lambda \int_{a(x)}^{x} K(x,s)\,\varphi(s)\,\mathrm{d}s + q\,f(x) \tag{B.1}$$

   – où p and q sont des constantes réelles données – forment une class qui n'est pas toujours en accord avec la théorie de VOLTERRA. En effet, quelquesunes d'entre elles ont plutôt un caractère fonctionnel et non intégral, en ce sens qu'elles se réduisent à des équations fonctionnelles proprement dites, de sorte que leurs ensembles des solutions dépendent parfois d'une fonction arbitraire, ce qui est contraire aux théorèmes d'unicité donnés par la théorie classique de VOLTERRA. Le plus simple exemple, possédant cette particularité, a été par C. POPOVICI, qui considère l'équation intégrale

$$\int_{-x}^{x} \phi(s)\,\mathrm{d}s = f(x). \tag{B.2}$$

   On voit aisément que, si $\varphi_1(x)$ en est une solution, l'équation précédente admet l'ensemble des solutions donné par

$$\varphi(x) = \varphi_1(x) + \Psi(x), \tag{B.3}$$

   où $\Psi(x)$ est une fonction impaire arbitraire, bornée et intégrable dans l'intervalle donné $(-a,a)$, dans lequel varie x. Si $f(x)$ est impaire, avec $f(0)=0$, l'équation intégrale (B.2) s'écrit sous la forme

$$\int_0^x \phi(s)\,\mathrm{d}s - \int_0^{-x} \phi(s)\,\mathrm{d}s = f(x) = \frac{f(x)}{2} - \frac{f(-x)}{2},$$

ou encore

$$\int_0^x \phi(s)\,\mathrm{d}s - \frac{f(x)}{2} = \int_0^{-x} \phi(s)\,\mathrm{d}s - \frac{f(-x)}{2},$$

qui montre que le premier membre est une fonction paire en x, $\Psi(x) = \Psi(-x)$,

$$\int_0^x \phi(s)\,\mathrm{d}s = \frac{f(x)}{2} + \Psi(x),$$

d'où, par dérivation, en supposant dérivables les fonctions $f(x)$ et $\Psi(x)$

$$\varphi(x) = \frac{f'(x)}{2} + \Psi_1(x),$$

qui représente l'ensemble (B.3) des solutions de (B.2), $\Psi_1(x)$ étant ce qu'est $\Psi(x)$ dans (B.3), c'est à dire, une fonction impaire, bornée et intégrable dans $(-a,a)$.

2. Arétons-nous un peu sur l'exemple précédent. L'équation intégrale (B.2) est ainsi résoluble si $f(x)$ est impaire et dérivable. Mais lorsque $f(x)$ n'est pas impaire ou dérivable, ou les deux à la fois, l'équation parait êgale à $f(x)$ pour $x > 0$ et à $-f(x)$ pour $x < 0$, $-a \leq x \leq a$: l'équation intégrale (B.2) sera remplacée par la suivante

$$\int_{-x}^x \varphi(s)\,\mathrm{d}s = F(x),$$

dont l'ensemble des solutions bornée et inégrables, sera donné par

$$\varphi(x) = \frac{F'(x)}{2} + \Psi_1(x),$$

ou, plus précisement, par

$$\varphi(x) = \begin{cases} \dfrac{f'(x)}{2} + \Psi_1(x) & \text{if } x > 0 \\ -\dfrac{f'(x)}{2} + \Psi_1(x) & \text{if } x < 0 \end{cases}$$

$\Psi_1(x)$ étant comme précédemment une fonction impaire arbitraire bornée et intégrable dans $(-a,a)$.

Annali di Mathematica Pura ed Applicata
Volume 54, Number 1, 33 – 36.

Integral Equations with Two Variable Limits[1]

by M. MICHEL GHERMANESCU (in Bucharest, Romania)

To M. Enrico Bompiani for his scientific jubilee

1. Integral equations with variable limits, discovered by VITO VOLTERRA at the end of the last century, constitute one of the powerful instruments of modern mathematical research, thanks to which the development of mathematical physics and other domains of research have made a considerable stride.

   Among these equations, those with two variable limits, comprised in the general linear type

$$p\,\varphi(x) = \lambda \int_{a(x)}^{x} K(x,s)\,\varphi(s)\,\mathrm{d}s + q\,f(x) \tag{B.4}$$

   – where p and q are given real constants – form a class which is not always in accord with the theory of VOLTERRA. Indeed, some of these have rather a functional and not integral character, in the sense that they reduce to equations properly called functional, of the kind that their sets of solutions depend sometimes on an arbitrary function, which is contrary to the uniqueness theorems given by the classical theory of VOLTERRA. The most simple example, possessing this particularity, has been given by C. POPOVICI, who considered the integral equation

$$\int_{-x}^{x} \phi(s)\,\mathrm{d}s = f(x). \tag{B.5}$$

   One easily sees that, if $\varphi_1(x)$ is a solution, the preceding equation admits the set of solutions given by

$$\varphi(x) = \varphi_1(x) + \Psi(x), \tag{B.6}$$

[1] Translated by Stephen M. Zemyan.

where $\Psi(x)$ is an arbitrary odd function, bounded and integrable in the given interval $(-a,a)$, in which x varies. If $f(x)$ is odd, with $f(0)=0$, the integral equation (B.5) is written under the form

$$\int_0^x \phi(s)\,\mathrm{d}s - \int_0^{-x} \phi(s)\,\mathrm{d}s = f(x) = \frac{f(x)}{2} - \frac{f(-x)}{2},$$

or again as

$$\int_0^x \phi(s)\,\mathrm{d}s - \frac{f(x)}{2} = \int_0^{-x} \phi(s)\,\mathrm{d}s - \frac{f(-x)}{2},$$

which shows that the left-hand side is an even function of x, $\Psi(x)=\Psi(-x)$,

$$\int_0^x \phi(s)\,\mathrm{d}s = \frac{f(x)}{2} + \Psi(x),$$

from which, by differentiation, supposing that the functions $f(x)$ and $\Psi(x)$ are differentiable,

$$\varphi(x) = \frac{f'(x)}{2} + \Psi_1(x),$$

which represents the set (B.6) of solutions of (B.5), $\Psi_1(x)$ being $\Psi(x)$ in (B.6), that is to say, an odd function, bounded and integrable in $(-a,a)$.

2. Let us reflect a little on the preceding example. The integral equation (B.2) is thus solvable if $f(x)$ is odd and differentiable. But when $f(x)$ is not odd or differentiable, or both, the equation appears to be unsolvable. We are going to show that in this case a *generalized solution* can be attributed to the integral equation (B.5).

   Indeed, we designate by $F(x)$ the odd, differentiable function, equal to $f(x)$ for $x>0$ and to $-f(x)$ for $x<0$, $-a\le x\le a$: The integral equation (B.5) will be replaced by the following

$$\int_{-x}^{x} \varphi(s)\,\mathrm{d}s = F(x),$$

whose set of bounded integrable solutions will be given by

$$\varphi(x) = \frac{F'(x)}{2} + \Psi_1(x),$$

or, more precisely, by

$$\varphi(x) = \begin{cases} \dfrac{f'(x)}{2} + \Psi_1(x) & \text{if } x>0 \\ -\dfrac{f'(x)}{2} + \Psi_1(x) & \text{if } x<0 \end{cases}$$

$\Psi_1(x)$ being, as before, an arbitrary odd function, bounded and integrable in $(-a,a)$.

B.2 Italian

B.2.1 Specialized Italian Vocabulary

English	**Italian**
adjoint	aggiunta
alternative	alternativa (*f*)
analytic	analytico
approximation theorem	teorema (*m*) d'approssimazione
bounded	limitato
bilinear form	forma (*f*) bilineare
boundary value problem	problema (*m*) con condizioni ai limiti
Cauchy sequence	successione (*f*) di Cauchy
Cauchy principal value	valore (*m*) principale secondo Cauchy
coefficient	coefficiente (*m*)
compact	compatto
complete	completo
complete orthonormal system	sistema (*m*) ortonormale completo
conjugate	coniugato
continuity	continuità (*f*)
continuous	continuo
contraction	contrazione (*f*)
contraction map	applicazione (*f*) restringente
convergence	convergenza (*f*)
convergent	convergente in media
cosine	coseno
cubic	cubico
determinant	determinante (*m*)
differentiable	derivabile
differential equation	equazione (*f*) differenziale
distance	distanza (*f*)
eigenfunction	autofunzione (*f*)
eigenvalue	autovalore (*m*)
eigenvector	autovettore (*m*)
finite	finito
first	primo
fixed point theorem	teorema (*m*) del punto fisso
formula	formula (*f*)
Fourier transformation	trasformazione (*f*) di Fourier

English	**Italian**
fourth	quarto
function	funzione (*f*)
gamma function	funzione (*f*) gamma
Hadamard's inequality	teorema (*m*) di Hadamard
Hermitian form	singolarità (*f*) eliminabile di Hermite
incomplete	incompleto
independent	independente
inequality	disuguaglianza (*f*)
infinite	infinito
infinite series	serie (*f*)
initial value problem	problema (*m*) a condizioni initiali
integrable	integrabile
integral equation	équation (*f*) intégrale
integration	integrazione (*f*)
integration by parts	integrazione (*f*) per parti
iterated	iterato
kernel	nucleo (*m*)
linear	lineare
linearly dependent	linearmente dipendente
Lipschitz condition	condizione (*f*) di Lipschitz
logarithmic	logaritmico
matrix	matrice (*f*)
method	metodo (*m*)
metric	metrica (*f*)
metric space	spazio (*m*) metrico
necessary	necessario
norm	norma (*f*)
normal	normale
orthogonal	ortogonale
orthogonalization	ortogonalizzione (*f*)
orthonormal system	sistema (*m*) ortonormale
point	punto (*m*)
polynomial	polinomio (*m*)
positive	positivo

English	Italian
positive definite	definito positivo
product kernel	nucleo prodotto (*m*)
proof	dimostrazione (*f*)
rectangle	rettangolo (*m*)
recursive	ricorrente
resolvent	risolvente (*f*)
Riemann integral	integrale (*m*) di Riemann
second	secondo
series	serie (*f*)
separable	separabile
sequence	sequenza (*f*)
singular	singolare
skew	antisimmetrica
solution	soluzione (*f*)
spectrum	spettro (*m*)
square	quadrato (*m*)
square integrable	di quadrato integrabile
substitution	sostituzione (*f*)
successive	successive
sufficient	sufficiente
sum	somma (*f*)
symmetric	simmetrico
term	termine (*m*)
third	terzo
transformation	trasformazione (*f*)
transpose matrix	matrice (*f*) trasposta
trapezoid rule	formula (*f*) dei trapezi
trigonometric	trigonometrico
vector	vettore (*m*)
zero	zero (*m*)
zero of a polynomial	zero (*m*) di un polinomio

B.2.2 Sample Translation from Italian into English

Annali di Mathematica Pura ed Applicata (4)
Volume **39** (1955), 229 – 244.

Equazioni integrali singolari del tipo di Carleman

FRANCESCO G. TRICOMI (a Torino).

A Mauro Picone nel suo 70mo compleanno.

Sunto. *Giustificazione rigorosa del metodo (euristico) di* CARLEMAN *per la risolutione di certe equazioni integrali contenenti il* valor principale *di un integrale (semplice), fondandosi sulla teoria della trasformazione finita di* HILBERT.

1. È ben noto che le equazioni integrali di FREDHOLM di seconda specie aventi nuclei della forma

$$K(x,y) = \frac{H(x,y)}{(y-x)^{\alpha}}$$

dove $H(x,y)$ è una funzione *sufficientemente regolare* (p. es. continua) ed α è un numero positivo *minore di uno*, non offrono particolari difficoltà perchè i successivi nuclei interati non contengono più, da un certo punto in poi, una potenza di $y-x$ al denominatore.

Molto più arduo è invece il caso $\alpha = 1$, che pur si presenta in importanti applicazioni, in cui bisogna considerare il *valor principale* (nel senso di CAUCHY) dell'integrale,[2] di guisa che l'equazione si presenta sotto l'aspetto

[2]Il valor principale (nel senso di CAUCHY) dell integrale di una funzione $f(x)$ dotata di un infinito (del primo ordine) nel punto c interno all'intervallo d'integrazione (a,b), valor principale che verrà indicato sovrapponendo un *asterisco* all'ordinario segno d'integrazione, si definisce notoriamente ponendo

$$\int_a^{*b} f(x)\,\mathrm{d}x = \lim_{\varepsilon\to 0}\int_a^{c-\varepsilon} + \int_{c+\varepsilon}^{b} f(x)\,\mathrm{d}x. \tag{B.7}$$

$$\varphi(x) - \lambda \int_a^{*b} \frac{H(x,y)}{(y-x)^\alpha}\, \varphi(y)\, \mathrm{d}y = F(x). \tag{B.8}$$

Meglio ancora – osservato che, sotto condizioni poco restrittive per H, il rapporto incrementale

$$H^*(x,y) = \frac{H(x,y) - H(x,x)}{y-x}$$

risulterà una funzione limitata o almeno sommabile (nel senso di LEBESGUE) – converrà scrivere l'equazione nel modo seguente

$$\varphi(x) - \lambda\, H(x,x) \int_a^{*b} \frac{\varphi(y)}{y-x}\, \mathrm{d}y = F(x) + \lambda \int_a^b H^*(x,y)\, \varphi(y)\, \mathrm{d}y.$$

Invero, si vede così che la cosa essenziale per dominare le equazioni del typo (B.8), è di saper risolvere un'equazione del tipo

$$\frac{\varphi(x)}{H(x,x)} - \lambda \int_a^{*b} \frac{\varphi(y)}{y-x}\, \mathrm{d}y = f(x)$$

ovvero

$$\boxed{a(x)\, \varphi(x) - \lambda \int_{-1}^{*1} \frac{\varphi(y)}{y-x}\, \mathrm{d}y = f(x)} \tag{B.9}$$

dove $a(x) = 1/H(x,x)$ e $f(x)$ sono due funzioni assegnate e, per comodità formale, si è supposto che l'intervallo fondamentale sia l'intervallo $(-1,1)$.

Le equazioni della forma (B.8) e, più particolare, l'equazione (B.9) canonica, possono con ragione dirsì, come qui faremo, *del tipo di Carleman* perchè risolute fin dal 1922 da quest'A. in un'elegante Memoria[3] rimasta ingiustamente semi-dimenticata per lunghi anni.

Noi esporremo più innanzi (Sect. 3) il metodo risolutivo di CARLEMAN o, più esattamente, il secondo e il migliore dei due methodi, tratti dalla teoria delle funzioni analitiche, usati nella Memoria succitata. Data però che non sembra facile giustificare i vari passagi in modo conforme alle moderne esigenze di rigore, il metodo di CARLEMAN verrà qui considerato da un punto vista puramente euristico e nei tre paragrafi che seguono (Sect. 4.6) la formula risolutiva a cui esso conduce verrà $\ll$ *legalizzata* $\gg$, delimitando un suo, abbastanza largo campo di validità, con metodi di genere diverso. Precede il §2 in cui vengono richiamati alcune proprietà della *trasformazione di Hilbert*, indispensabili per l'accennata legittimazione delle formule di CARLEMAN.

[3] T. CARLEMAN, *Sur la résolution de certaines équations intégrales*, Arkiv for Mat. Astron. och Fysik, **16** (1922), 19 pp.

Annali di Mathematica Pura ed Applicata (4)
Volume **39** (1955), 229 – 244.

Singular Integral Equations of Carleman Type[4]

FRANCESCO G. TRICOMI (in Torino).

To Mauro Picone on his 70th birthday.

Abstract. *Rigorous justification of the (heuristic) method of* CARLEMAN *for the resolution of certain integral equations containing the* principal value *of an integral (simple), basing it on the theory of the finite* HILBERT *transformation.*

1. It is well known that FREDHOLM integral equations of the second kind having kernels of the form

$$K(x,y) = \frac{H(x,y)}{(y-x)^{\alpha}}$$

where $H(x,y)$ is a *sufficiently regular* (for example continuous) function and α is a positive number *less than one*, do not present any particular difficulties because the successive iterated kernels no longer contain, from a certain point onwards, a power of $y-x$ in the denominator.

Much more difficult is instead the case $\alpha = 1$, which arises in important applications, in which it is necessary to consider the *principal value* (in the sense of CAUCHY) of the integral,[5] in such a way that the equation appears as

[4]Translated by Stephen M. Zemyan.

[5]The principal value (in the sense of CAUCHY) of the integral of a function $f(x)$ with a given infinite value (of the first order) at the point c interior to the interval of integration (a,b) (the principal value will be indicated by putting an *asterisk* on the ordinary sign of integration) is defined notationally by putting

$$\int_a^{*b} f(x)\,\mathrm{d}x = \lim_{\varepsilon\to 0}\int_a^{c-\varepsilon} + \int_{c+\varepsilon}^{b} f(x)\,\mathrm{d}x. \tag{B.10}$$

$$\varphi(x) - \lambda \int_a^{*b} \frac{H(x,y)}{(y-x)^{\alpha}} \varphi(y)\,\mathrm{d}y = F(x). \tag{B.11}$$

Better yet—it can be observed that, under less restrictive conditions on H, the difference relation

$$H^*(x,y) = \frac{H(x,y) - H(x,x)}{y-x}$$

will result in a bounded function or at least summable (in the LEBESGUE sense) – it enables us to write the equation in the following manner

$$\varphi(x) - \lambda\, H(x,x) \int_a^{*b} \frac{\varphi(y)}{y-x}\,\mathrm{d}y = F(x) + \lambda \int_a^b H^*(x,y)\,\varphi(y)\,\mathrm{d}y.$$

Indeed, it is seen that the essential thing to master the equations of type (B.11) is to know how to solve an equation of the type

$$\frac{\varphi(x)}{H(x,x)} - \lambda \int_a^{*b} \frac{\varphi(y)}{y-x}\,\mathrm{d}y = f(x)$$

or

$$\boxed{a(x)\,\varphi(x) - \lambda \int_{-1}^{*1} \frac{\varphi(y)}{y-x}\,\mathrm{d}y = f(x)} \tag{B.12}$$

where $a(x) = 1/H(x,x)$ and $f(x)$ are two prescribed functions and, for formal convenience, it is supposed that the fundamental interval is the interval $(-1,1)$.

Equations of the form (B.11) and, more particularly, the canonical equation (B.12), can be rightly called, as we have done, *of Carleman type* because it was solved by him in 1922 in an elegant Memoir[6] that remained unjustly half-forgotten for many years.

We previously (Sect. 3) established CARLEMAN'S method of solution, or, more exactly, the second and the better of the two methods, treated with the theory of analytic functions, used in the Memoir cited above. However, given that it does not seem easy to justify the various passages in a manner that conforms to the modern exigencies of rigor, the method of CARLEMAN will be considered here from a purely heuristic point of view and in the three paragraphs which follow (Sect. 4.6) the solution to which it leads will be ≪ *legitimized* ≫, in a large enough field of validity, with methods of a different kind. Section 2 precedes it, in which some properties of the *Hilbert transformation*, indispensible for the legitimization of CARLEMAN'S formulas, are restated.

[6] T. CARLEMAN, *Sur la résolution de certaines équations intégrales*, Arkiv for Mat. Astron. och Fysik, **16** (1922), 19 pp.

B.3 German

B.3.1 Specialized German Vocabulary

English	German
adjoint	adjungierte
alternative	Alternative (*f*)
analytic	analytisch
approximation theorem	Approximationssatz (*m*)
bounded	beschränkt
bilinear form	Bilinearform (*f*)
boundary value problem	Randwertproblem (*n*)
Cauchy sequence	CauchyFolge (*f*)
Cauchy principal value	Cauchyscher Hauptwert (*n*)
coefficient	Koeffizient (*m*)
compact	kompakt
complete	vollständig
complete orthonormal system	vollständiges Orthonormalsystem (*n*)
conjugate	konjugiert
continuity	Stetigkeit (*f*)
continuous	stetig
contraction	Verjüngung (*f*)
contraction map	kontrahierende Abbildung (*f*)
convergence	Konvergenz (*f*)
convergent	konvergent
cosine	Cosinus (*m*)
cubic	kubisch
determinant	Determinante (*f*)
differentiable	differenzierbar
differential equation	Differentialgleichung (*f*)
distance	Abstand (*m*)
eigenfunction	Eigenfunktion (*f*)
eigenvalue	Eigenwert (*m*)
eigenvector	Eigenvektor (*m*)
finite	endlich
first	erste
fixed point theorem	Fixpunktsatz (*m*)
formula	Formel (*f*)
Fourier transformation	Fouriersche Integraltransformation (*f*)

English	German
fourth	vierte
function	Funktion (*f*)
gamma function	Gammafunktion (*f*)
Hadamard's inequality	Hadamardscher Determinantensatz (*m*)
Hermitian	hermitesche
incomplete	unvollständig
independent	unabhängig
inequality	Ungleichung (*f*)
infinite	unendlich
infinite series	unendliche Reihe (*f*)
initial value problem	Anfangswertproblem (*n*)
integrable	integrierbar
integral equation	Integralgleichung (*f*)
integration	Integration (*f*)
integration by parts	partielle Integration (*f*)
iterated	iteriert
kernel	Kern (*m*)
linear	linear
linearly dependent	linear abhängig
Lipschitz condition	Lipschitz-Bedingung
logarithmic	logarithmisch
matrix	Matrix (*f*)
method	Methode (*f*)
metric	Metrik (*f*)
metric space	metrischer Raum
necessary	notwendig
norm	Norm (*f*)
normal	Normale (*f*)
orthogonal	orthogonal, senkrecht
orthogonalization	Orthogonalisierung (*f*)
orthonormal system	Orthonormalsystem (*n*)
point	Punkt (*m*)
polynomial	Polynom (*n*)
positive	positiv

English	German
positive definite	positiv definit
product kernel	Produktkern (*m*)
proof	Beweis (*m*)
rectangle	Rechteck (*n*)
recursive	rekursiv
resolvent	Resolvente (*f*)
Riemann integral	Riemannsches Integral (*n*)
second	zweite
series	Reihe (*f*)
separable	separabel
sequence	Folge (*f*)
singular	singulär
skew	schief
solution	Lösung (*f*)
spectrum	Spektrum (*n*)
square	Quadrat (*n*)
square integrable	quadratintegrabel
substitution	Substitution (*f*)
successive	sukzessiv
sufficient	hinreichend
sum	Summe
symmetric	symmetrisch
term	Term (*m*)
third	dritte
transformation	Transformation (*f*)
transpose matrix	transponierte Matrix (*f*)
trapezoid rule	Trapezregel (*f*)
trigonometric	trigonometrisch
vector	Vektor (*m*)
zero	Null (*f*)
zero of a polynomial	Nullstelle (*f*) eines Polynoms

B.3.2 Sample Translation from German into English

DINGHAS, A.
Math. Zeitschr. Bd. 70, S. 174–189 (1958)

Zur Existenz von Fixpunkten bei Abbildungen vom Abel-Liouvilleschen Typus

Von

ALEXANDER DINGHAS

1. Einleitung. Es sei $f(x,y)$ eine im Rechteck

$$R\colon 0 \le x \le a, \quad |y| \le b \quad (0 < a,b < +\infty) \tag{B.13}$$

eindeutige reele stetige Funktion von x und y, und es bezeichne C die Gesamtheit aller eindeutigen, reellen, stetigen Funktionen $g(x)$, $|g(x)| \le b$, die im Intervall $J\colon 0 \le x \le a$ definiert sind und im Nullpunkt verschwinden. Ist dann μ eine reelle Zahl aus dem Intervall $(0,1]$, so liefert die Transformation

$$T_\mu(g) = \frac{1}{\Gamma(\mu)} \int_0^x (x-t)^{\mu-1} f(t,g(t))\,\mathrm{d}t^{7} \tag{B.14}$$

eine Abbildung von C. Die Frage nach der Existenz eines Fixpunktes, d.h. einer Funktion $y(x)$, $y(x) \in C$ mit der Eigenschaft

$$y(x) = \frac{1}{\Gamma(\mu)} \int_0^x (x-t)^{\mu-1} f(t,y(t))\,\mathrm{d}t \tag{B.15}$$

für alle $x \in J_1$ ($J_1 = [0,a_1]$) mit einem geeigneten a_1, $0 < a_1 \le a$ ist bekanntlich äquivalent mit der Auffindung einer (stetigen) Lösung der Integrodifferentialgleichung

$$\frac{\mathrm{d}}{\mathrm{d}x}\left\{\frac{1}{\Gamma(1-\mu)} \int_0^x (x-t)^{-\mu} y(t)\,\mathrm{d}t\right\} = f(x,y) \tag{B.16}$$

---

[7] Sowohl in diesem, als auch in den nachfolgenden Integralen soll der Wert der Integrale für $x = 0$ durch Stetigkeit definiert werden. Es gilt dann

$$\Gamma(\mu+1) \lim_{x\to 0} \{T_\mu(g(x))/x^\mu\} = f(0,0),$$

und somit ist $T_\mu(g(x)) = O(x^\mu)$.

mit $y(0) = 0$. Ist $0 < \mu < 1$, so kann man diese Tatache durch Multiplikation von (B.15) mit $(\tau - x)^{-\mu}$ und Integration von $x = 0$ bis $x = \tau$ unter Heranziehung einer einfachen Transformationsformel, die in **2.** ausführlich beweisen wird, Leicht nachweisen. Ebenso leicht kann gezeigt werden, daß aus (B.16) ohne weiteres (B.15) folgt. Läßt man in (B.14) $\mu \to 1$ konvergieren, so konvergiert $T_\mu(g)$ gegen $\int_0^x f(t,g)\,\mathrm{d}t$. Man kann entsprechend zeigen, daß (bei festem g) der Ausdruck

$$\frac{1}{\Gamma(1-\mu)} \int_0^x (x-t)^{-\mu} g(t)\,\mathrm{d}t$$

für $\mu \to 1$ gegen $f(x)$ konvergiert. Man setze in der Tat

$$D_\mu(x) = \frac{1}{\Gamma(1-\mu)} \int_0^x (x-t)^{-\mu} g(t)\,\mathrm{d}t - \frac{g(x)}{\Gamma(2-\mu)}. \tag{B.17}$$

Dann ist

$$D_\mu(x) = \frac{1}{\Gamma(1-\mu)} \int_0^x (x-t)^{-\mu} \{g(t) - g(x)\}\,\mathrm{d}t. \tag{B.18}$$

Man wähle bei vorgegebenem $\varepsilon > 0$ und einem $x > 0$ x_1 $(0 < x_1 < x)$ so, daß im Intervall $x_1 \le t \le x$ die Ungleichung

$$|g(t) - g(x)| \le \varepsilon$$

gilt. Dann wird

$$\left|D_\mu(x)\right| \le \frac{1}{\Gamma(1-\mu)} \int_0^{x_1} (x-t)^{-\mu} |g(t) - g(x)|\,\mathrm{d}t + \frac{\varepsilon}{\Gamma(2-\mu)},$$

und somit

$$\limsup_{\mu \to 1} \left|D_\mu(x)\right| \le \varepsilon.$$

Das beweist die Behauptung.

Diese Zusammenhänge würden zweifellos keine selbständige Abhandlung rechtfertigen, wenn nicht zugleich zwei weitere Gründe hinzukämen, die mich veranlaßt haben, diese Note zu publizieren. Erstens die Tatsache, daß hier dem Weierstraßschen Approximationssatz sowohl beim Beweis der Eindeutigkeit der Lösung als auch beim Konvergenzbeweis der Approximationsfolge eine wesent-liche Rolle zukommt und zweitens, daß die hier gegebenen Bedingungen nicht ohne weiteres gelockert werden können. Es herrschen also sowohl in dem Fall von **3.** und **4.**, als auch in demjenigen von **6.** und **7.** dieselben Verhältnisse, wie in den Fällen $\mu = 1$ von ROSENBLATT-NAGUMO-PERRON-HAVILAND und KRASNOSELSKI-KREIN-LUXENBURG, für welche man weiß, daß die zugehörigen Lipschitz-Bedingungen nicht abgeschwächt werden können, ohne daß man die Eindeutigkeit der Lösung und die Konvergenz der Iterationsfolge einbüßt.

DINGHAS, A.
Math. Zeitschr. Bd. 70, S. 174–189 (1958)

On the Existence of Fixed Points of Mappings of Abel–Liouville Type[8]

By

ALEXANDER DINGHAS

1. Introduction. Let $f(x,y)$ be a single-valued real continuous function of x and y in the rectangle

$$R\colon\ 0 \le x \le a, \quad |y| \le b \quad (0 < a, b < +\infty) \tag{B.19}$$

and let C denote the collection of all single-valued, real, continuous functions $g(x)$, $|g(x)| \le b$, defined in the interval $J\colon\ 0 \le x \le a$ which vanish at zero. If μ is a real number in the interval $(0,1]$, then the transformation

$$T_\mu(g) = \frac{1}{\Gamma(\mu)} \int_0^x (x-t)^{\mu-1} f(t, g(t))\, \mathrm{d}t^{9} \tag{B.20}$$

yields a mapping of C. The question of the existence of fixed points, i.e., of a function $y(x)$, $y(x) \in C$ with the property that

$$y(x) = \frac{1}{\Gamma(\mu)} \int_0^x (x-t)^{\mu-1} f(t, y(t))\, \mathrm{d}t \tag{B.21}$$

for all $x \in J_1$ ($J_1 = [0, a_1]$) with suitable a_1 ($0 < a_1 \le a$) is well known to be equivalent to the existence of a (continuous) solution to the integrodifferential equation

$$\frac{\mathrm{d}}{\mathrm{d}x} \left\{ \frac{1}{\Gamma(1-\mu)} \int_0^x (x-t)^{-\mu} y(t)\, \mathrm{d}t \right\} = f(x,y) \tag{B.22}$$

with $y(0) = 0$. If $0 < \mu < 1$, then one can easily prove these properties by multiplication of (B.21) by $(\tau - x)^{-\mu}$ and integration from $x = 0$ to $x = \tau$ drawing

---

[8] Translated by Stephen M. Zemyan.

[9] As usual, the value of the following integrals at $x = 0$ will be defined by continuity. Then we have

$$\Gamma(\mu+1) \lim_{x \to 0} \left\{ T_\mu(g(x)) / x^\mu \right\} = f(0,0),$$

and consequently $T_\mu(g(x)) = O(x^\mu)$.

upon a simple transformation formula, as we will prove in detail in **2.** It is just as easy to show without further work that (B.21) follows from (B.22). If we let $\mu \to 1$ in (B.20), then $T_\mu(g)$ converges to $\int_0^x f(t,g)\,\mathrm{d}t$. It can be correspondingly shown (with fixed g) that the expression

$$\frac{1}{\Gamma(1-\mu)} \int_0^x (x-t)^{-\mu} g(t)\,\mathrm{d}t$$

converges to $g(x)$ as $\mu \to 1$. Indeed, one can set

$$D_\mu(x) = \frac{1}{\Gamma(1-\mu)} \int_0^x (x-t)^{-\mu} g(t)\,\mathrm{d}t - \frac{g(x)}{\Gamma(2-\mu)}. \tag{B.23}$$

Then, we have

$$D_\mu(x) = \frac{1}{\Gamma(1-\mu)} \int_0^x (x-t)^{-\mu} \{g(t) - g(x)\}\,\mathrm{d}t. \tag{B.24}$$

One can choose in advance an $\varepsilon > 0$ and an $x_1 > 0$ $(0 < x_1 < x)$ so that the inequality

$$|g(t) - g(x)| \le \varepsilon$$

holds in the interval $x_1 \le t \le x$. Then we will have

$$\left|D_\mu(x)\right| \le \frac{1}{\Gamma(1-\mu)} \int_0^{x_1} (x-t)^{-\mu} |g(t) - g(x)|\,\mathrm{d}t + \frac{\varepsilon}{\Gamma(2-\mu)},$$

and consequently

$$\limsup_{\mu \to 1} \left|D_\mu(x)\right| \le \varepsilon.$$

The assertion is proven.

Undoubtedly, any independent discussion would justify this relationship, as I have done by publishing this note, if two further facts would be added at the same time. First, the properties that play an essential role in the Weierstrass approximation theorem together with the uniqueness proof associated with the convergence proof of the approximation sequence, and second, that the condition given here cannot be further weakened. This relation holds also in the case of 3. and 4., as well as in that of 6. and 7., as in the case $\mu = 1$ of ROSENBLATT-NAGUMO-PERRON-HAVILAND and KRASNOSELSKI-KREIN-LUXENBURG, for which it is known that the associated Lipschitz condition cannot be further weakened, without which one loses the uniqueness of the solution and the convergence of the iteration sequence.

References

[Da] Davis, H.T.: Introduction to Nonlinear Differential and Integral Equations. Dover Publications, New York (1962)

[Er] Erdogan, F.: Approximate solutions of systems of singular integral equations. SIAM J. Appl. Math. **17**(6), 1041–1059 (1969)

[Fr] Fredholm, I.: Sur une classe d'equations fonctionelles. Acta Math. **27**(1), 365–390 (1903)

[Ko] Kondo, J.: Integral Equations. Clarendon Press/Oxford University Press, Oxford (1992)

[Kr] Krasnov, M.L., Kiselev, A.I., Makarenko, G.I.: Equazioni Integrali. Traduzione de Alexandr Machov, Distribuito da Libreria Italia-URSS, Piazza della Repubblica 47, Roma (1983)

[Li] Linz, P.: Analytical and Numerical Methods for Volterra Equations, Studies in Applied Mathematics. SIAM, Philadelphia (1985)

[Mi] Mikhlin, S.G.: Integral Equations and their Applications to Certain Problems in Mechanics, Mathematical Physics and Technology, 2nd revised edn. Macmillan, New York (1964)

[Mo] Moiseiwitsch, B.L.: Integral Equations. Dover Publications, New York (2005)

[Mu] Muskhelishvili, N.I.: Singular Integral Equations. Boundary Problems of Function Theory and their Application to Mathematical Physics. Noordhoff International Publishing, Leyden (1953)

[PS] Porter, D., Stirling, D.: Integral Equations, a Practical Treatment, From Spectral Theory to Applications. Cambridge Texts in Applied Mathematics, Cambridge University Press, Cambridge (1990)

[Pv] Privalov, I.I.: Introduction to the Theory of Functions of a Complex Variable, 8th edn. State Technical Publishers, Gosudarstv. Izdat. Tehn.-Ten. Lit., Moscow-Leningrad (1948)

[Sa] Saaty, T.L.: Modern Nonlinear Equations. Dover Publications, New York (1981)

[Tri] Tricomi, F.G.: Integral Equations. Dover Publications, New York (1985)

[We] Weisstein, E.W.: CRC Concise Encyclopedia of Mathematics. 2nd edn. Chapman and Hall/CRC, Boca Raton (2003)

[Wi] Widder, D.V.: The Laplace Transform. Princeton University Press, Princeton (1946)

[Yo] Yosida, K.: Lectures on Differential and Integral Equations. Dover Publications, New York (1991)

S.M. Zemyan, *The Classical Theory of Integral Equations: A Concise Treatment*,
DOI 10.1007/978-0-8176-8349-8, © Springer Science+Business Media, LLC 2012

Additional Reading

[Ba] Baker, C.T.H.: The Numerical Treatment of Integral Equations. Oxford University Press, New York (1977)

[Bo] Bôcher, M.: An Introduction to the Study of Integral Equations. Cambridge Tracts in Mathematics and Mathematical Physics, No. 10. Hafner Publishing Company, New York (1971). Reprint of the 2nd edition, 1914

[Bu] Burton, T.A.: Volterra Integral and Differential Equations, 2nd edn. Mathematics in Science and Engineering, vol. 202. Elsevier, Amsterdam (2005)

[Ca1] Carleman, T.: La théorie des équations intégrales singulières et ses applications. Ann. l'I. H. P. **tome 1**(4), 401–430 (1930)

[Ca2] Carleman, T.: Zur Theorie der linearen Integralgleichungen. Mathematische Zeitschrift **9**(3–4), 196–217 (1921)

[Coc] Cochran, J.A.: The Analysis of Linear Integral Equations. McGraw-Hill, New York (1972)

[Cor1] Corduneanu, C.: Differential and Integral Equations. Chelsea Publishing Company, New York (1977)

[Cor2] Corduneanu, C.: Integral Equations and Applications. Cambridge University Press, Cambridge (1991)

[CH] Courant, R., Hilbert, D.: Methods of Mathematical Physics, vol. 1, Ch. III. Interscience, New York (1964)

[De] Delves, L.M., Mohamed, J.L.: Computational Methods for Integral Equations. Cambridge University Press, Cambridge (1985)

[G] Garabedian, P.R.: Partial Differential Equations, Ch. 10. AMS Chelsea Publishing, Providence (1998)

[GR] Gradshteyn, I.S., Ryzhik, I.M.: Table of Integrals, Series, and Products, Corrected and Enlarged Edition. Academic Press, London (1980)

[GL] Guenther, R.B., Lee, J.W.: Partial Differential Equations of Mathematical Physics and Integral Equations, Ch. 7. Prentice-Hall, Englewood Cliffs (1988)

[Ha] Hammerstein, A.: Nichtlineare Integralgleichungen nebst Anwendungen. Acta Math. **54**, 117–176 (1930)

[Hi] Hilbert, D.: Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen. Teubner, Leipzig (1912). Reprinted by Chelsea Publishing Company, New York, 1953

[Hoc] Hochstadt, H.: Integral Equations. Pure and Applied Mathematics. Wiley-Interscience Publications, New York (1989)

[Hoh] Hoheisel, G.: Integral Equations. Ungar Publishing Company, New York (1968)

[Je] Jerry, A.J.: Introduction to Integral Equations with Applications, 2nd edn. Wiley-Interscience Publications, New York (1999)

[Ka] Kanwal, R.P.: Linear Integral Equations, 2nd edn. Birkhauser Boston, Boston (1997)

[Ki] Kincaid, D., Cheney, W.: Numerical Analysis: Mathematics of Scientific Computing, 3rd edn. American Mathematical Society, Providence, Rhode Island (2002)

[LR] Lakshmikantham, V., Rao, M.R.M.: Theory of Integro-differential Equations, Stability and Control: Theory, Methods and Applications, vol. 1. Gordon and Breach Science Publishers, Lausanne (1995)

[Ob] Oberhettinger, F.: Tables of Mellin Transforms. Springer, New York (1974)

[Pi] Pipkin, A.C.: A Course on Integral Equations, Texts in Applied Mathematics, vol. 9. Springer, New York (1991)

[PM] Polyanin, A.D., Manzhirov, A.V.: Handbook of Integral Equations, 2nd edn. Chapman and Hall/CRC Press, Boca Raton (2008)

[Pr] Prössdorf, S.: Linear Integral Equations, Analysis IV. Springer, Berlin. Encyclopaedia Math. Sci. **27**, 1–125 (1991)

[Sm] Smithies, F.: Integral Equations, Cambridge Tracts in Mathematics and Mathematical Physics, No. 49. Cambridge University Press, London (1958)

[Ti] Titchmarsh, E.C.: Introduction to the Theory of Fourier Integrals, 3rd edn. Chelsea Publishing Co., New York (1986)

[Tr2] Tricomi, F.G.: On the finite Hilbert transformation. Quart. J. Math. (Oxford) **2**, 199–211 (1951)

[Trj] Trjitzinsky, W.F.: General theory of singular integral equations with real kernels. Trans. Am. Math. Soc. **46**, 202–279 (1939)

[Vo] Volterra, V.: Theory of Functionals and of Integral and Integro-differential Equations. Dover Publications, New York (1959)

[Za] Zabreyko, P.P., Koshelev, A.I., Krasnosel'skii, M.A., Mikhlin, S.G., Rakovshchik, L.S., Stet'senko, V.Ya.: Integral Equations – A Reference Text. Noordhoff International Publishing, Leyden (1975)

Index

A
Abel's equation, 267
  generalized, 271
  second kind, 272
adjoint equation, 2, 14, 20
airfoil equation, 275

B
Banach Fixed Point Theorem, 214
Bessel function
  first kind, 175, 196
    order 0, 194
  modified, 166
  second kind
    order 0, 194
Bessel's Inequality, 93
boundary value problem, 184
bounded set, 89

C
Carleman's formula, 283
Cauchy criterion, 32
coefficients
  Fourier, 93
compact, 89
complete orthonormal system, 93
construction
  resolvent kernel, 60
Contraction
  Principle of, 214
contraction mapping, 213
contraction operator, 213
convergence
  in the mean, 92
  uniform, 32
  weak, 97
convolution
  finite Hilbert, 275
  Fourier, 249
  Laplace, 152
  Mellin, 248
criterion
  Riesz-Fischer, 92
cubic splines, 80

D
differential equation
  Airy, 203
  Bessel
    order ν, 184, 194
  Euler, 184
  Hermite, 204
  linear, 183
  linear system, 187
Dini's Theorem, 131
Dirac δ-function, 178
discretization error, 181
distance
  mean square, 92

E
eigenfunction, 19, 52
  of a kernel, 8
eigenfunctions, 252

S.M. Zemyan, *The Classical Theory of Integral Equations: A Concise Treatment*,
DOI 10.1007/978-0-8176-8349-8, © Springer Science+Business Media, LLC 2012

eigenspace, 3, 18
eigenvalue
  algebraic multiplicity, 3, 18
  geometric multiplicity, 3, 18, 55
  of a kernel, 6, 15, 51
  of a matrix, 3
  regular, 7
eigenvector, 3
expansion
  finite bilinear, 110
exponential order, 152

F
fixed point, 213
Fourier series, 94
Fredholm determinant, 15
Fredholm equation
  linear
    first kind, 118, 133
    second kind, 1, 13
  nonlinear
    second kind, 211
Fredholm theorem, 48
  Alternative, 23, 58
  First, 11, 17, 53
  Second, 11, 21, 56
  Third, 11, 23, 58
  Fourth, 11, 23, 54
free term, 1, 14, 151
function
  analytic, 33
  continuous vector, 288
  entire, 33
  holomorphic, 33
  meromorphic, 33

G
Guide to Computation
  Collocation, 76
  Fredholm equations, 24

H
Hammerstein's Theorem, 219
Hilbert's Formula, 124
Hilbert-Schmidt Theorem, 120

I
identity
  binomial coefficient, 176
  combinatorial, 175
incomplete, 93
inequality
  Cauchy-Schwarz, 32
inhomogeneous, 14
initial value problem, 184, 197
inner product, 3
integral
  beta, 246
  Cauchy principal value, 245
  improper
    convergent, 244
    divergent, 244
integrating derivatives, 185
integrodifferential equation, 205

K
kernel
  Cauchy, 273
  complex symmetric, 95
  convolution, 166
  eigenvalues, 252
  Gaussian, 266
  Hermitian, 85, 95
  Hilbert, 273
  iterated, 34, 154
  logarithmic, 281
  positive, 127
  positive definite, 127
  real symmetric, 95
  resolvent, 7, 16, 35, 125, 155
  separable, 1
  singular, 243, 267
  skew Hermitian, 106
  trace, 64, 98
  triangular, 219
  Volterra, 153

L
Lipschitz condition, 215, 273

M
matrix
  adjoint, 4
  fundamental, 187
  inverse, 2
  norm, 288
Mercer's Theorem, 129, 194
metric, 213
  complete, 213
metric space, 213

mutually orthogonal, 86
mutually orthonormal, 86

N
nodes, 78
norm, 31
  matrix, 288
  vector function, 288
numerical instability, 182
numerical methods, 69
  collocation, 73
  quadrature, 78, 180
  Rayleigh and Ritz, 142
  successive approximation, 153
  traces, 140
  uniform approximation, 69

O
operator
  bounded, 89
  completely continuous, 97
  conjugate transpose, 90
  continuous, 96
  Fredholm, 88
  Hermitian, 96
  norm, 89
  normal, 96
ordinary point, 184
orthogonal, 3
Orthonormalization
  Gram-Schmidt, 87

P
Parseval's Identity, 93
Parseval's relation, 250
piecewise continuous, 152
polynomial
  characteristic, 3
  Chebychev, 285
  Legendre, 86, 120

Q
$Q(a,b)$, 1

R
regular, 16, 51
residual, 73
Riemann integrable, 85
Riemann zeta function, 148

S
series
  Neumann, 35, 155
  resolvent, 157
singular equation, 243
  first kind, 252, 253, 256, 277
  Fox, 254
  Lalesco-Picard, 260
  second kind, 252, 253, 256, 262, 263, 276
singular point
  regular, 184
solution, 35
successive approximation, 37, 217
successive substitution, 33, 302
system
  coupled, 297
  Fredholm, 300
    first kind, 300
    second kind, 300
    third kind, 300
  linear, 287
  singular, 307
    first kind, 307
    second kind, 307
    third kind, 307
  Volterra, 290

T
$T(a,b)$, 151
theorem
  Dini, 131
  fixed point, 214
  Fourier integral, 248
  Hammerstein, 219
  Hilbert-Schmidt, 121
  Mercer, 129
  successive approximation, 41, 157
  successive substitution, 36
trace, 98
Trace Lemma, 98
transform
  finite Hilbert, 275
  Fourier, 248
  Fourier cosine, 251, 259
  Fourier sine, 251
  Gaussian, 266
  Hilbert, 276
  inverse Fourier cosine, 251

transform (*cont.*)
inverse Fourier sine, 251
inverse Mellin, 247
Laplace, 152, 290
Mellin, 247

V
vector space
complex, 3
Volterra equation, 197
linear
first kind, 172
second kind, 151
third kind, 174
nonlinear
first kind, 235
second kind, 211

W
weights, 78